

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 15:22:33 ; Search time 165 Seconds
(without alignments)
274.248 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 625
Sequence: 1 ECGGTIMGNATVFCPEYGP.....FQWKDLSIKGVYEACMAL 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	100.0	117	5	ABB06244
2	625	100.0	589	5	AAE13606
3	625	100.0	3461	5	ABB05008
4	625	100.0	3461	5	ABB57065
5	625	100.0	3461	7	ADBE60142
6	572	91.5	3427	7	ADJ69818
7	572	91.5	3460	7	ABBO5007
8	572	91.5	3460	7	ADBE60143
9	572	91.5	3460	7	ADN95182
10	572	91.5	3460	8	ADQ20919
11	572	91.5	3460	8	ADQ88204
12	572	91.5	3470	4	ABG25297
13	499	79.8	432	5	AAE13605
14	499	79.8	623	5	AAE13607
15	224	35.8	84	8	ABO56700
16	150	30.4	84	8	ABG55221
17	107	17.1	28	4	AAW74141
18	107	17.1	28	4	AAW61384
19	107	17.1	28	4	ABG55918
20	107	17.1	28	5	ABG44060
21	83	13.3	639	1	AAE40212
22	81.5	13.0	468	4	ABG25295
23	81	13.0	616	2	AAW55977
24	81	13.0	616	2	AAW55976
25	81	13.0	616	2	AAW55980

26	81	13.0	616	2	AAW55979	AAW55979	Aspergill
27	81	13.0	616	3	AAE15181	AAE15181	Aspergill
28	81	13.0	616	3	AAE15176	AAE15176	Aspergill
29	81	13.0	616	3	AAE15178	AAE15178	Aspergill
30	81	13.0	616	3	AAE15184	AAE15184	Aspergill
31	81	13.0	616	3	AAE15180	AAE15180	Aspergill
32	81	13.0	616	3	AAE15179	AAE15179	Aspergill
33	81	13.0	621	2	AAW55978	AAW55978	Aspergill
34	81	13.0	621	3	AAE15182	AAE15182	Aspergill
35	81	13.0	621	3	AAE15183	AAE15183	Aspergill
36	81	13.0	630	1	AAE1876	AAE1876	Sequence
37	81	13.0	640	4	AAE17741	AAE17741	A. niger
38	81	13.0	640	4	AAE161905	AAE161905	A. niger
39	74.5	11.9	834	4	ABG25296	ABG25296	Novel hum
40	73	11.7	1095	6	ABP96630	ABP96630	Alpha-amyl
41	71.5	11.4	91	8	ABO55728	ABO55728	Human gen
42	70.5	11.3	271	5	ABE29970	ABE29970	Streptococ
43	70.5	11.3	279	5	ABE25985	ABE25985	Streptococ
44	69.5	11.1	667	4	ABE70456	ABE70456	Drosophila
45	68.5	11.0	326	7	ADH70051	ADH70051	C. neoform
46	68.5	11.0	572	8	ADJ50225	ADJ50225	Oil-asso
47	68.5	11.0	611	8	ADJ49056	ADJ49056	Oil-asso
48	68.5	11.0	611	8	ADJ59749	ADJ59749	B. subtil
49	68	10.9	426	6	ABP96582	ABP96582	Cow IgE h
50	68	10.9	459	3	AAE20638	AAE20638	Arabidops
51	68	10.9	461	3	AAE20637	AAE20637	Arabidops
52	68	10.9	462	8	ADK47072	ADK47072	Streptococ
53	68	10.9	472	6	ABU02501	ABU02501	S. pneumo
54	68	10.9	473	8	ADK94638	ADK94638	Novel S.
55	68	10.9	512	3	AAE20636	AAE20636	Arabidops
56	68	10.9	553	7	ADL17839	ADL17839	Chimeric
57	68	10.9	574	7	ADBE4863	ADBE4863	Human pro
58	67.5	10.8	3623	2	AAE10197	AAE10197	Rat cubil
59	67	10.7	90	3	AAE10192	AAE10192	Arabidops
60	67	10.7	110	3	AAE15092	AAE15092	Arabidops
61	67	10.7	222	5	AAE15172	AAE15172	Human sec
62	67	10.7	222	5	ABG65027	ABG65027	Human alb
63	67	10.7	222	8	ADL78294	ADL78294	Albunin f
64	67	10.7	225	6	ABR41331	ABR41331	Human DIT
65	67	10.7	267	5	AAE13449	AAE13449	Human sec
66	67	10.7	267	5	ABG65030	ABG65030	Human alb
67	67	10.7	267	8	ADL78297	ADL78297	Albunin f
68	67	10.7	339	4	AAU09964	AAU09964	Human cyt
69	67	10.7	356	7	ABO65401	ABO65401	Klebsiell
70	67	10.7	489	5	AAE13171	AAE13171	Human sec
71	67	10.7	489	5	ABG65029	ABG65029	Human alb
72	67	10.7	489	8	ADL78296	ADL78296	Albunin f
73	67	10.7	514	5	AAE48334	AAE48334	Human spe
74	67	10.7	528	5	AAE13399	AAE13399	Human sec
75	67	10.7	743	7	ADG87098	ADG87098	Glucosamyl
76	66.5	10.6	89	2	AAE137665	AAE137665	Amino acyl
77	66.5	10.6	544	6	ABU25791	ABU25791	Protein e
78	66	10.6	624	4	AAE48170	AAE48170	Thermosac
79	66	10.6	743	7	ADG87094	ADG87094	Glucosamyl
80	66	10.6	1098	2	AAE13392	AAE13392	Hemophil
81	66	10.6	1524	4	ABG25288	ABG25288	Novel hum
82	65.5	10.5	53	4	AAE41552	AAE41552	Propionib
83	65.5	10.5	126	4	ABM38041	ABM38041	Propionib
84	65.5	10.5	126	4	ABM78909	ABM78909	C. glutam
85	65.5	10.5	209	3	AAE51026	AAE51026	Arabidops
86	65.5	10.5	234	8	ADG86177	ADG86177	Arabidops
87	65.5	10.5	269	4	ABE69533	ABE69533	Drosophila
88	65.5	10.5	372	8	ADG23806	ADG23806	Bacterial
89	65.5	10.5	374	3	AAE41650	AAE41650	Arabidops
90	65.5	10.5	465	6	ABJ25380	ABJ25380	Arabidops
91	65.5	10.5	549	6	ABG90782	ABG90782	C. glutam
92	65.5	10.5	549	8	ADG47934	ADG47934	Beta-vuln
93	65.5	10.5	564	3	AAE41649	AAE41649	Arabidops
94	65.5	10.5	577	3	AAE41648	AAE41648	Arabidops
95	65.5	10.5	611	8	ADJ50073	ADJ50073	Oil-asso
96	65.5	10.5	620	7	ADW47277	ADW47277	Protocodac
97	65.5	10.5	624	6	ABJ25980	ABJ25980	Aspergill
98	65.5	10.5	941	5	ABE93761	ABE93761	Herbicida

99 65.5 10.5 1042 4 ABB60137
100 65.5 10.5 1102 4 ABB62029

AbB60137 Drosophila
AbB62029 Drosophila

ALIGNMENTS

RESULT 1
ID ABB06244 standard; protein; 117 AA.

XX ABB06244;

XX 21-MAY-2002 (first entry)

XX Mouse reelin protein CR-50 epitope region protein SEQ ID NO:2.

XX Mouse; reelin protein CR-50 epitope region; elucidation; neuron;
XX cerebral disturbance; reelin protein; neuroprotective.

XX Mus musculus.

XX JP2002017361-A.

XX 22-JAN-2002.

XX 04-JUL-2000; 2000JP-00202801.

XX 04-JUL-2000; 2000JP-00202801.

XX (RIKE) RIKEN KK.

XX WPI; 2002-221707/28.

XX N-PSDB; ABL40165.

XX Reelin protein CR-50 epitope region, useful for diagnosis and treatment

XX of cerebral disturbance.

XX Claim 3; Page 11-12; 16pp; Japanese.

XX The present sequence represents the mouse reelin protein CR-50 epitope
XX region, which contains the CR-50 antibody recognition site and is free
XX from F-spondin domains and repetitive sites. Also described are: (1) an
XX expression vector comprising a polynucleotide encoding a reelin protein
XX epitope region; (2) host cells with transfected the expression vector;
XX (3) polypeptides prepared by culture of the host cells; and (4)
XX polynucleotides comprising the 351 base sequence given in ABL40165 which
XX encodes the 117 amino acid sequence given in ABB06244; and (5) use of the
XX polynucleotide for diagnosis and/or treatment of diseases caused by
XX abnormal positioning of neural cells, and stimulation of association of
XX reelin protein. The mouse reelin protein CR-50 epitope region has
XX neuroprotective activity, and can be used in the diagnosis and treatment
XX of cerebral disturbance due to an abnormal reelin gene and positioning of
XX neurons

XX Sequence 117 AA;

Query Match 100.0%; Score 625; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.7e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ECGGTIMGNNAVTFCEPYGPRLTTTCINTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 1 ECGGTIMGNNAVTFCEPYGPRLTTTCINTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
QY 61 NNTADMWIOLEKIRAPSNVSTVTHILYLPBEAKGSVQFQWKDLSRVGEVYEACMAL 117
DB 61 NNTADMWIOLEKIRAPSNVSTVTHILYLPBEAKGSVQFQWKDLSRVGEVYEACMAL 117

RESULT 2
ID AAE13606 standard; protein; 589 AA.

XX AAE13606;
XX 26-FEB-2002 (first entry)

XX Mus musculus truncated reelin protein.

XX Mouse; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
XX polymicrogyria; ectopic gray matter.

XX Mus musculus.

XX Key Location/Qualifiers

XX Peptide 1..27

XX Protein /label= Signal_peptide
XX 28..589
XX /label= Mature_truncated_reelin_protein

XX EP1149844-A2.

XX 31-OCT-2001.

XX 11-APR-2001; 2001EP-00303411.

XX 11-APR-2000; 2000JP-00109954.

XX (RIKE) RIKEN KK.

XX Mikoshiba K, Tabata H, Nakajima K;

XX WPI; 2002-019320/03.

XX N-PSDB; AAD22754.

XX Novel truncated Reelin protein containing F-spondin domain and CR-50
XX recognition site of Reelin protein, but not having Reelin repeat site,
XX useful to treat diseases including agyria due to abnormal neuron
XX alignment.

XX Claim 4; Page 26-30; 47pp; English.

XX The invention relates to a truncated Reelin protein comprising a F-
XX spondin domain and a CR-50 recognition site but no reelin protein repeat
XX site. Reelin is an essential molecule in developing a normal laminated
XX structure of cerebrum. The truncated reelin protein and its DNA are
XX useful for treating diseases including agyria, polymicrogyria, and
XX ectopic gray matter due to abnormal neuronal alignment. Truncated reelin
XX protein DNA is useful in gene therapy. The present sequence is Mus
XX musculus truncated reelin protein

XX Sequence 589 AA;

Query Match 100.0%; Score 625; DB 5; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.6e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ECGGTIMGNNAVTFCEPYGPRLTTTCINTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 230 ECGGTIMGNNAVTFCEPYGPRLTTTCINTTASVLOFSIGSGCRFSYSDPSITVSYAK 289
QY 61 NNTADMWIOLEKIRAPSNVSTVTHILYLPBEAKGSVQFQWKDLSRVGEVYEACMAL 117
DB 230 NNTADMWIOLEKIRAPSNVSTVTHILYLPBEAKGSVQFQWKDLSRVGEVYEACMAL 346

RESULT 3
ID ABB05008 standard; protein; 3461 AA.
XX ABB05008;
XX 21-MAR-2002 (first entry)
XX Mouse reelin protein SEQ ID NO:2.

XX Mouse; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
 KW extracellular glycoprotein; nootropic; antilipemic; Alzheimer's disease;
 KW neurodegenerative disorder; neuronal regeneration; cognitive function;
 KW lipid metabolism disease; memory; developmental disorder.
 XX Mus musculus.
 OS
 XX US6323177-B1.
 XX
 XX 27-NOV-2001.
 XX
 XX 16-JUN-1999; 99US-00334220.
 XX
 XX 16-JUN-1999; 99US-00334220.
 XX
 XX (SUJN-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Curran T, D'Arcangelo G;
 PI
 DR MPI: 2002-096596/13.
 XX
 DR N-PSDB; ABA92603.
 XX
 XX Novel composition useful for screening compounds that modulate reelin
 PT binding to low density lipoprotein receptor, comprising an isolated
 PT Reelin polypeptide and low density lipoprotein receptor.
 PT
 PS Claim 16; Col 47-64; 45pp; English.
 XX
 XX The present invention describes a composition (I) comprising an isolated
 CC reelin protein (II) bound to an isolated low density lipoprotein receptor
 CC (LDLR) (III). (II) is an extracellular glycoprotein of approximately 385
 CC kDa containing a small region of similarity with F-glycophorin at the N
 CC terminus, a stretch of positively charged amino acids at the C terminus,
 CC and a series of eight internal repeats of 350-390 amino acids, each
 CC repeat containing two related sub-domains that flank a pattern of
 CC conserved cysteine residues known as an epidermal growth factor (EGF)-
 CC like motif. (I) has neuroprotective, nootropic and antilipemic
 CC activities, and can be used as a modulator of reelin-LDLR interaction.
 CC (I) is useful in screen for compounds that modulate reelin binding to an
 CC LDLR, in an assay system, where the assay system comprises a microplate
 CC array and an automated robotic microprocessor controlled system for
 CC adding and removing reagents to the microplate array. The compounds
 CC identified by the above screening method are useful as therapeutic agents
 CC to provide or alleviate a diverse spectrum of diseases including
 CC neurodegenerative disorders such as Alzheimer's disease, to facilitate
 CC neuronal regeneration after injury, to prevent or alleviate lipid
 CC metabolism diseases, to enhance cognitive functions and memory or to
 CC ameliorate other developmental disorders. The present sequence represents
 CC mouse (Mus musculus) reelin, which is used in the exemplification of the
 CC present invention
 XX
 XX Sequence 3461 AA;
 SQ
 Query Match 100.0%; Score 625; DB 5; Length 3461;
 Best Local Similarity 100.0%; Pred. No. 1.9e-62;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOCGTMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGSCFSSYSDPSITVSVAK 60
 DB 230 EOCGTMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGSCFSSYSDPSITVSVAK 289
 QY 61 NNTADMIOLEKTRAPSNVSTVHIILYLPBEAKGESVQFQKODSLVGEVYECWAL 117
 DB 290 NNTADMIOLEKTRAPSNVSTVHIILYLPBEAKGESVQFQKODSLVGEVYECWAL 346
 RESULT 4
 ID ABB57065 standard; protein, 3461 AA.
 XX ABB57065;
 AC
 XX

DT 07-MAR-2002 (first entry)
 DE Mouse ischemic condition related protein sequence SEQ ID NO:129.
 XX
 XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischemic condition; ischaemic disease.
 XX
 XX Mus musculus.
 OS
 XX WO200108188-A2.
 XX
 XX 22-NOV-2001.
 XX
 XX 18-MAY-2001; 2001WO-JP004192.
 XX
 XX 18-MAY-2001; 2000JP-00145977.
 XX
 XX (UYN-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 PI
 DR MPI: 2002-034733/04.
 XX
 DR N-PSDB; ABI99284.
 XX
 XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.
 PT
 PS Claim 2; Page 385-400; 2690pp; English.
 XX
 XX The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention
 XX
 XX Sequence 3461 AA;
 SQ
 Query Match 100.0%; Score 625; DB 5; Length 3461;
 Best Local Similarity 100.0%; Pred. No. 1.9e-62;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOCGTMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGSCFSSYSDPSITVSVAK 60
 DB 230 EOCGTMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGSCFSSYSDPSITVSVAK 289
 QY 61 NNTADMIOLEKTRAPSNVSTVHIILYLPBEAKGESVQFQKODSLVGEVYECWAL 117
 DB 290 NNTADMIOLEKTRAPSNVSTVHIILYLPBEAKGESVQFQKODSLVGEVYECWAL 346
 RESULT 5
 ID ADE60142 standard; protein, 3461 AA.
 XX ADE60142;
 AC
 XX 29-JAN-2004 (first entry)
 DT
 XX Rat Protein NP_035391, SBQ ID NO 6041.
 DE
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GENO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX MPI: 2003-268312/26.
XX DR GENBANK; NF_035391.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1, page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3461 AA:
SQ
Query Match 100.0%; Score 625; DB 7; Length 3461;
Best Local Similarity 100.0%; Pred. No. 1.9e-62;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EOCGIMGNANTFCEPYGPRELITTCNTTASVLOFSIGSGCRFSYSPSITVSYAK 60
DB 230 EOCGIMGNANTFCEPYGPRELITTCNTTASVLOFSIGSGCRFSYSPSITVSYAK 289
QY 61 NNTADMIOLEKIRAPSNVSTVIHILYLPEAKGESVQFQKODSLRVEVEYACAL 117
DB 230 NNTADMIOLEKIRAPSNVSTVIHILYLPEAKGESVQFQKODSLRVEVEYACAL 346

AC ADJ69818;
XX
XX 06-MAY-2004 (first entry)
XX
XX Human heat mitochondrial protein as a therapeutic target SegID1624.
XX
XX mitochondrial; human; screening assay; diabetes mellitus;
XX Huntington's disease; osteoarthritis;
XX Leber's hereditary optic neuropathy; LHON;
XX mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
XX osteopathic; ophthalmological; cyostatic.
XX
XX Homo sapiens.
XX
XX WO2003087768-A2.
XX
XX 23-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010870.
XX
XX 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-0389987P.
XX PR 20-SEP-2002; 2002US-0412418P.
XX
XX (MITO-) MITOKOR.
XX (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX Warnock DE;
XX MPI: 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
XX PT treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX PT with the disease.
XX
XX Claim 1; SEQ ID NO 1624; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
XX CC for therapeutic intervention in treating a disease associated with
XX CC altered mitochondrial function. Specifically, it refers to a method for
XX CC identifying proteins of the human heart mitochondrial proteome that are
XX CC useful for drug screening assays, as well as therapeutic targets. The
XX CC present invention describes a method for identifying such proteins that
XX CC can be used in the treatment of various diseases associated with altered
XX CC mitochondrial function including diabetes mellitus, Huntington's disease,
XX CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX CC compositions have neuroprotective, nootropic, antidiabetic,
XX CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX CC cyostatic activities. This polypeptide sequence is a human heart
XX CC mitochondrial protein of the invention.
XX
XX Sequence 3427 AA:
SQ
Query Match 91.5%; Score 572; DB 7; Length 3427;
Best Local Similarity 90.6%; Pred. No. 2.8e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 EOCGIMGNANTFCEPYGPRELITTCNTTASVLOFSIGSGCRFSYSPSITVSYAK 60
DB 229 EOCGIMGNANTFCEPYGPRELITTCNTTASVLOFSIGSGCRFSYSPSITVSYAK 288
QY 61 NNTADMIOLEKIRAPSNVSTVIHILYLPEAKGESVQFQKODSLRVEVEYACAL 117
DB 289 NNSADMIOLEKIRAPSNVSTVIHILYLPEAKGENVQFQKODSLRVEVEYACAL 345

RESULT 7

AB05007
ID ABB05007 standard; protein; 3460 AA.
XX
AC ABB05007;
XX
DT 21-MAR-2002 (first entry)
XX
DE Human reelin protein SEQ ID NO:1.
XX
KW Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; nootropic; antilipemic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder.
XX
OS Homo sapiens.
XX
PN US623177-B1.
XX
PD 27-NOV-2001.
XX
PF 16-JUN-1999; 99US-00334220.
XX
PR 16-JUN-1999; 99US-00334220.
XX
PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Curran T, D'arcangelo G;
XX
DR WPI; 2002-096596/13.
XX
DR N-PSDB; ABA92604.
XX
PT Novel composition useful for screening compounds that modulate Reelin
PT binding to low density lipoprotein receptor, comprising an isolated
PT Reelin polypeptide and low density lipoprotein receptor.
XX
PS Claim 16; Col 31-48; 45pp; English.
XX
CC The present invention describes a composition (I) comprising an isolated
CC reelin protein (II) bound to an isolated low density lipoprotein receptor
CC (LDLR) (III). (II) is an extracellular glycoprotein of approximately 385
CC kDa containing a small region of similarity with F-pondin at the N
CC terminus, a stretch of positively charged amino acids at the C terminus,
CC and a series of eight internal repeats of 350-390 amino acids, each
CC repeat containing two related sub-domains that flank a pattern of
CC conserved cysteine residues known as an epidermal growth factor (EGF)-
CC like motif. (I) has neuroprotective, nootropic and antilipemic
CC activities, and can be used as a modulator of reelin-LDLR interaction.
CC (I) is useful in screen for compounds that modulate reelin binding to an
CC LDLR, in an assay system, where the assay system comprises a microplate
CC array and an automated robotic microprocessor controlled system for
CC adding and removing reagents to the microplate array. The compounds
CC identified by the above screening method are useful as therapeutic agents
CC to provide or alleviate a diverse spectrum of diseases including
CC neurodegenerative disorders such as Alzheimer's disease, to facilitate
CC neuronal regeneration after injury, to prevent or alleviate lipid
CC metabolism diseases, to enhance cognitive functions and memory or to
CC ameliorate other developmental disorders. The present sequence represents
CC human reelin, which is used in the exemplification of the present
CC invention
XX
SQ Sequence 3460 AA;
XX
Query Match 91.5%; Score 572; DB 5; Length 3460;
Best Local Similarity 90.6%; Pred. No. 2.8e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
XX
QY 1 EOCCTIMHGNVAFCEPYGPRELTTTCCLNTTASVLOFSIGSSGCFPSYSDPITSYAK 60
DB 229 EOCGALMHGNAVFCEPYGPRELTTTCCLNTTASVLOFSIGSSGCFPSYSDPITSYAK 268
XX
QY 61 NNTADMIOLEKIRAPSNVSTVTHIILYLPBEAKGESVVOFQWKQDSLNGEYEAQMAL 117
DB 289 NNSADWIOLEKIRAPSNVSTVTHIILYLPBEAKGESVVOFQWKQDSLNGEYEAQMAL 345
XX

RESULT 8
ADE60143
ID ADE60143 standard; protein; 3460 AA.
XX
AC ADE60143;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P78509, SEQ ID NO 6043.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'ureo D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
XX
DR GENBANK; P78509.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPRO at
CC ftp.wipro.int/pub/published_pct_sequences.
XX
SQ Sequence 3460 AA;
XX
Query Match 91.5%; Score 572; DB 7; Length 3460;
Best Local Similarity 90.6%; Pred. No. 2.8e-56;
XX

Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMGNATVFCPEPYGPRELTTTCINTTTASVLOFSIGSGSCRFSPSPSITVSXAK 60
DB 229 EOCGAIMGNATVFCPEPYGPRELTTGINTTTASVLOFSIGSGSCRFSPSPSITVLVXAK 288

QY 61 NNTADWIOLEKTRAPSNVSTVIHILYLPBEAKGESVQPOKODSLRVGEVTEACMAL 117
DB 289 NNSADMIOLEKTRAPSNVSTIIHILYLPEDAKGENVQPOKODENLRVGEVTEACMAL 345

RESULT 9
ADN95182
ID ADN95182 standard; protein; 3460 AA.
AC ADN95182;
XX
XX
DT 01-JUL-2004 (first entry)
XX
DE Human BRC/LBC-related protein sequence SegID104.
XX
XX growth; differentiation; blood endothelial cell; BRC;
XX lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
XX lymphatic growth agent; VEGF-D; angiogenic; cytostatic;
XX vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
XX inflammatory disease; cancer metastasis; lymphatic system; human.
XX
XX Homo sapiens.
XX
XX WO2003080640-A1.
XX
XX 02-OCT-2003.
XX
XX 07-MAR-2003; 2003WO-US006900.
XX
XX 07-MAR-2002; 2002US-0363019P.
XX
XX (LUDM-) LUDWIG INST CANCER RES.
XX (LICN) LICENTIA LTD.
XX
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX MPI: 2003-876899/81.
XX DR N-PSDB; ADN95183.
XX
XX Example 1; SEQ ID NO 104; 176pp; English.

This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprising contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGFR-3; and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an angiogenic, cytostatic, vasotropic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note:

CC This sequence does not appear in the specification but was obtained by
CC the indexer using the source data given in table 14 of the specification.
XX
XX Sequence 3460 AA;
SQ

QY 1 EOCGTIMGNATVFCPEPYGPRELTTTCINTTTASVLOFSIGSGSCRFSPSPSITVSXAK 60
DB 229 EOCGAIMGNATVFCPEPYGPRELTTGINTTTASVLOFSIGSGSCRFSPSPSITVLVXAK 288

QY 61 NNTADWIOLEKTRAPSNVSTVIHILYLPBEAKGESVQPOKODSLRVGEVTEACMAL 117
DB 289 NNSADMIOLEKTRAPSNVSTIIHILYLPEDAKGENVQPOKODENLRVGEVTEACMAL 345

RESULT 10
ADQ20919
ID ADQ20919 standard; protein; 3460 AA.
AC ADQ20919;
XX
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 3739.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
XX Homo sapiens.
XX
XX WO2004048938-A2.
XX
XX 10-JUN-2004.
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
XX 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX PA Aziz N, Ginsburg WM, Zlocznik A;
XX PI WPI: 2004-441208/41.
XX
XX Example 2; SEQ ID NO 3739; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX Sequence 3460 AA;
SQ

Query Match 91.5%; Score 572; DB 8; Length 3460;
Best Local Similarity 90.6%; Pred. No. 2.8e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 ECGTTMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 60
 DB 229 ECGGAIMHGNAVTFCEPYGPRELTTTGLNTTASVLOFSIGSGSCRFPSYSDPSITVLYAK 288
 QY 61 NNTADMIQLEKIRAPSNVSTVTHILYLPBAGKESVQFQMKDLSLVGEVYECMAL 117
 DB 289 NNSADMIQLEKIRAPSNVSTIHLIYLPBDAKENVQFQMKDNLKRVGEVYECMAL 345

RESULT 11
 ADQ88204
 ID ADQ88204 standard; protein; 3460 AA.
 XX
 AC ADQ88204;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human 3533 protein, a reelin serine protease protein RELN.
 XX
 KW human; cardiovascular disorder; thrombotic disorder;
 KW differential expression; gene therapy; aberrant vascularisation;
 KW atherosclerosis; thrombosis; coronary artery disease; hyperlipidemia;
 KW dyslipidaemia; high blood pressure; heart failure; cardiac;
 KW thrombolytic; anticoagulant; antilipemic; hypotensive; cardiac;
 KW reelin serine protease; RELN; EC 3.4.21; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN WO2004063340-A2.
 XX
 PD 29-JUL-2004.
 XX
 PP 13-JAN-2004; 2004WO-US000393.
 XX
 PR 13-JAN-2003; 2003US-0439683P.
 PR 05-FEB-2003; 2003US-0445216P.
 PR 18-FEB-2003; 2003US-0448036P.
 PR 12-MAR-2003; 2003US-0454189P.
 PR 25-MAR-2003; 2003US-0457541P.
 PR 29-APR-2003; 2003US-0466411P.
 PR 08-MAY-2003; 2003US-0469041P.
 PR 10-JUN-2003; 2003US-0477414P.
 PR 13-JUN-2003; 2003US-0478560P.
 PR 24-JUL-2003; 2003US-0489772P.
 PR 28-JUL-2003; 2003US-0490660P.
 PR 03-SEP-2003; 2003US-049838P.
 PR 22-SEP-2003; 2003US-0504786P.
 PR 24-SEP-2003; 2003US-0505570P.
 PR 17-OCT-2003; 2003US-0512418P.
 PR 27-OCT-2003; 2003US-0514660P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Stagliano NE, Healy A, Acton SL, Galvin KM, Donoghue MA;
 PI Rogrigue-Way A, Tomlinson JE;
 PI
 DR N-PSDB; ADQ88203.
 DR
 DR WPI; 2004-553729/53.
 PT
 PT identifying a compound for treating a cardiovascular or thrombotic
 PT disorder by combining a compound to be tested with e.g., a 9380, 9462,
 PT 8701 or 2419 polypeptide or with a host cell expressing the polypeptide
 PT and detecting the binding.
 XX
 PS Claim 1; SEQ ID NO 46; 512pp; English.
 XX
 CC This invention relates to a novel compound that is capable of treating a
 CC cardiovascular or thrombotic disorder. Specifically, it refers to the
 CC identification of nucleic acid molecules, and the encoded proteins
 CC thereof, which are differentially expressed in cardiovascular disease
 CC states relative to their normal expression in non-diseased tissue. The
 CC present invention describes test compounds (i.e. small molecules,
 CC peptides or antibodies) that can bind to and modulate the activity of

CC these differentially expressed membrane-bound polypeptides, where binding
 CC is detected by a competition binding assay, immunoassay or yeast two-
 CC hybrid assay. Accordingly, pharmaceutical compositions can be developed
 CC and used via gene therapy to treat aberrant vascularisation,
 CC atherosclerosis, thrombosis, coronary artery disease, hyperlipidemia,
 CC dyslipidaemia, high blood pressure or heart failure. As such, they
 CC exhibit cardiac, thrombolytic, anticoagulant, antilipemic, hypotensive
 CC and cardiac activities. This polypeptide sequence is a human protein
 CC that is differentially expressed in a patient with a cardiovascular
 CC disorder, given in an exemplification of the invention.

XX
 SQ Sequence 3460 AA;
 XX

Query Match 91.5%; Score 572; DB 8; Length 3460;
 Best Local Similarity 90.6%; Pred. No. 2.8e-56;
 Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 ECGTTMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 60
 DB 229 ECGGAIMHGNAVTFCEPYGPRELTTTGLNTTASVLOFSIGSGSCRFPSYSDPSITVLYAK 288
 QY 61 NNTADMIQLEKIRAPSNVSTVTHILYLPBAGKESVQFQMKDLSLVGEVYECMAL 117
 DB 289 NNSADMIQLEKIRAPSNVSTIHLIYLPBDAKENVQFQMKDNLKRVGEVYECMAL 345

RESULT 12
 ABG25297
 ID ABG25297 standard; protein; 3470 AA.
 XX
 AC ABG25297;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #25288.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PP 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 PI
 DR N-PSDB; AAG89484.
 DR
 DR WPI; 2001-639362/73.
 PT
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 55656; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AB00010-AB030377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 3470 AA;
SQ

Query Match 91.5%; Score 572; DB 4; Length 3470;
Best Local Similarity 90.6%; Pred. No. 2.8e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELITTCNTTTASVLOFSISGSGCRFSSDPSITVSXK 60
DB 239 EOCGAIMGNNAVTFCEPYGPRELITTCNTTTASVLOFSISGSGCRFSSDPSITVSXK 298
61 NNTADWIOLEKIRAPSNVSTVTHILYLPBEAKGESVOPQMKDSLAVGEVYBACWAL 117
299 NNSADWIOLEKIRAPSNVSTVTHILYLPBEAKGESVOPQMKDSLAVGEVYBACWAL 355

RESULT 13
AAE13605
ID AAE13605 standard; protein; 432 AA.
XX
AC AAE13605;
XX
DT 26-FEB-2002 (first entry)
XX
DE Xenopus laevis truncated reelin protein.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
XX polymicrogyria; ectopic gray matter.
XX
OS Xenopus laevis.
XX

Key Location/Qualifiers
FT Peptide 1..26
FT /label= Signal_peptide
FT Protein 27..432
FT /label= Mature_truncated_reelin_protein

EP1149844-A2.
XX
PN 31-OCT-2001.
XX
PD 11-APR-2001; 2001EP-00303411.
XX
PF 11-APR-2000; 2000JP-00109954.
XX
PR (RIKE) RIKEN KK.
XX
PA Mikoshiba K, Tabata H, Nakajima K;
XX
PI WPI; 2002-019320/03.
XX
DR N-PSDB; AAD22753.
XX

Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment.
XX
PS Claim 3; Page 16-19; 47p; English.
XX
CC The invention relates to a truncated Reelin protein comprising a F-
spondin domain and a CR-50 recognition site but no reelin protein repeat

CC site. Reelin is an essential molecule in developing a normal laminated
CC structure of cerebrum. The truncated reelin protein and its DNA are
CC useful for treating diseases including agyria, polymicrogyria, and
CC ectopic gray matter due to abnormal neuronal alignment. Truncated reelin
CC protein DNA is useful in gene therapy. The present sequence is Xenopus
XX laevis truncated reelin protein
XX

Sequence 432 AA;
SQ

Query Match 79.8%; Score 499; DB 5; Length 432;
Best Local Similarity 74.4%; Pred. No. 5.1e-49;
Matches 87; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELITTCNTTTASVLOFSISGSGCRFSSDPSITVSXK 60
DB 231 EOCGAIMGNNAVTFCEPYGPRELITTCNTTTASVLOFSISGSGCRFSSDPSITVSXK 290
61 NNTADWIOLEKIRAPSNVSTVTHILYLPBEAKGESVOPQMKDSLAVGEVYBACWAL 117
291 NNSSMWPLEIRAPSNVSTVTHILYLPBEAKGESVOPQMKDSLAVGEVYBACWAL 347

RESULT 14
AAE13607
ID AAE13607 standard; protein; 623 AA.
XX
AC AAE13607;
XX
DT 26-FEB-2002 (first entry)
XX
DE Xenopus laevis intact reelin protein.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
XX polymicrogyria; ectopic gray matter.
XX
OS Xenopus laevis.
XX

Key Location/Qualifiers
FT Misc-difference 484
FT /note= "Encoded by CGA"
FT Misc-difference 495
FT /note= "Encoded by CAT"

EP1149844-A2.
XX
PN 31-OCT-2001.
XX
PD 11-APR-2001; 2001EP-00303411.
XX
PF 11-APR-2000; 2000JP-00109954.
XX
PR (RIKE) RIKEN KK.
XX
PA Mikoshiba K, Tabata H, Nakajima K;
XX
PI WPI; 2002-019320/03.
XX
DR N-PSDB; AAD22779.
XX

Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment.
XX
PS Example 1; Fig 1A; 47p; English.
XX
CC The invention relates to a truncated Reelin protein comprising a F-
spondin domain and a CR-50 recognition site but no reelin protein repeat
CC site. Reelin is an essential molecule in developing a normal laminated
CC structure of cerebrum. The truncated reelin protein and its DNA are
CC useful for treating diseases including agyria, polymicrogyria, and
CC ectopic gray matter due to abnormal neuronal alignment. Truncated reelin
CC protein DNA is useful in gene therapy. The present sequence is Xenopus
XX laevis intact reelin protein

XX Sequence 623 AA;
SQ
Query Match 79.8%; Score 499; DB 5; Length 623;
Best Local Similarity 74.4%; Pred. No. 8.5e-49;
Matches 87; Conservative 17; Mismatches 13; Indels 0; Gaps 0;
QY 1 ECGCTMHGNAVFCEPYPGRELTTCTANTTASVLOFSIGSGSCRPYSDPSTVSYAK 60
DB 231 ECGALMHGNAVTFCDPYPGRELITVQMTTASVLOFSIGSGSCRPYSDDPGIVSYTK 230
QY 61 NNTADNITQLEKIRAPSNVSTVHILYLPBEAKGESVQFQWKDSLVRGEVYEAACWAL 117
DB 291 NNSGSMWPLERISAPSNVSTIHIYLPBEAKGENVFRKQENMQAGDYEAACWAL 347
RESULT 15
ABO56700
ID ABO56700 standard; protein; 84 AA.
XX
AC ABO56700;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #2934.
XX
KM Human; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANK/) HANKEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI, 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 30334; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above, an
CC method of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing

CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030194704
XX
XX
SQ Sequence 84 AA;
QY Query Match 35.8%; Score 224; DB 8; Length 84;
DB Best Local Similarity 88.9%; Pred. No. 6.1e-18;
Matches 40; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 73 RAPSNTVTHILYLPBEAKGESVQFQWKDSLVRGEVYEAACWAL 117
DB 3 RAPSNTVTHILYLPBEAKGENVQFQWKDSLVRGEVYEAACWAL 47

Search completed: August 4, 2005, 15:32:36
Job time : 169 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 15:27:08 ; Search time 43 Seconds
(Without alignments)
203.115 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 625
Sequence: 1 EOCGRIMHNAVTFCPEYGP.....PQWKQDSLRYGEVYACMAL 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents MA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_5/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	100.0	3461	3	US-09-334-220-2
2	572	91.5	3460	3	US-09-334-220-1
3	81	13.0	616	4	US-09-236-063-1
4	81	13.0	640	3	US-09-351-814-13
5	70.5	11.3	568	4	US-09-248-796A-14918
6	68	10.9	462	4	US-09-583-110-3587
7	68	10.9	473	4	US-09-107-433-3273
8	68	10.9	884	4	US-09-902-540-13902
9	67.5	10.8	3623	4	US-09-341-461-2
10	67	10.7	339	4	US-09-802-371-2
11	67	10.7	356	4	US-09-489-039A-11918
12	66.5	10.6	569	4	US-09-514-245-22
13	66	10.6	658	1	US-08-409-995-5
14	66	10.6	658	3	US-08-685-467-5
15	66	10.6	658	3	US-08-913-942-5
16	66	10.6	658	4	US-09-684-707-5
17	66	10.6	1098	1	US-08-409-995-2
18	66	10.6	1098	1	US-08-685-467-2
19	66	10.6	1098	3	US-09-377-155-32
20	66	10.6	1098	3	US-08-913-942-2
21	66	10.6	1098	3	US-09-669-974-32
22	66	10.6	1098	3	US-09-258-347-44
23	66	10.6	1098	4	US-09-797-862-32
24	66	10.6	1098	4	US-09-684-707-2
25	65.5	10.5	126	4	US-09-603-208A-78
26	65.5	10.5	159	4	US-09-270-767-34992
27	65.5	10.5	159	4	US-09-270-767-50209

28	65.5	10.5	431	4	US-09-270-767-42950	Sequence 42950, A
29	65.5	10.5	540	4	US-09-538-092-793	Sequence 793, App
30	65.5	10.5	549	3	US-09-291-922-30	Sequence 30, App1
31	65	10.4	1187	1	US-08-201-697-2	Sequence 2, App1
32	64.5	10.3	213	2	US-08-716-284-2	Sequence 2, App1
33	64	10.2	533	4	US-09-302-540-13075	Sequence 13075, A
34	64	10.2	574	3	US-08-732-433-1	Sequence 1, App1
35	64	10.2	574	3	US-08-945-574-2	Sequence 2, App1
36	64	10.2	574	4	US-09-663-547B-2	Sequence 2, App1
37	63.5	10.2	280	4	US-09-248-796A-11297	Sequence 17297, A
38	63.5	10.2	337	2	US-08-923-856-4	Sequence 4, App1
39	63.5	10.2	337	3	US-09-216-294-4	Sequence 4, App1
40	63	10.1	148	4	US-09-252-991A-24972	Sequence 24972, A
41	63	10.1	382	1	US-08-049-282B-2	Sequence 2, App1
42	63	10.1	382	2	US-08-537-715-2	Sequence 2, App1
43	63	10.1	382	5	PCT-US94-04173-2	Sequence 2, App1
44	63	10.1	383	5	US-08-537-715-4	Sequence 4, App1
45	63	10.1	383	5	PCT-US94-04173-4	Sequence 4, App1
46	63	10.1	3913	4	US-09-949-016-10933	Sequence 10933, A
47	63	10.1	4377	4	US-09-949-016-6978	Sequence 6978, Ap
48	62.5	10.0	336	4	US-09-248-796A-17023	Sequence 17023, A
49	62.5	10.0	884	6	5208144-8	Patent No. 5208144
50	62.5	10.0	884	6	5208144-8	Patent No. 5208144
51	61.5	9.8	247	4	US-09-134-000C-6617	Sequence 6617, Ap
52	61.5	9.8	838	4	US-09-270-767-46665	Sequence 46665, A
53	61.5	9.8	1411	4	US-09-538-092-413	Sequence 413, App
54	61	9.8	375	4	US-08-978-404B-8	Sequence 8, App1
55	61	9.8	408	4	US-09-248-796A-16704	Sequence 16704, A
56	61	9.8	420	4	US-09-543-681A-6574	Sequence 6574, Ap
57	61	9.8	484	4	US-09-489-039A-8961	Sequence 8961, Ap
58	61	9.8	589	4	US-09-328-352-7592	Sequence 7592, Ap
59	61	9.8	1188	1	US-08-201-697-4	Sequence 4, App1
60	61	9.8	1246	4	US-09-949-016-8051	Sequence 8051, Ap
61	61	9.8	1246	4	US-09-949-016-8052	Sequence 8052, Ap
62	61	9.8	1246	4	US-09-949-016-8053	Sequence 8053, Ap
63	61	9.8	1246	4	US-09-949-016-8054	Sequence 8054, Ap
64	61	9.8	1246	4	US-09-949-016-8055	Sequence 8055, Ap
65	61	9.8	1274	4	US-09-949-016-8828	Sequence 8828, Ap
66	61	9.8	1274	4	US-09-949-016-8829	Sequence 8829, Ap
67	61	9.8	1274	4	US-09-949-016-8830	Sequence 8830, Ap
68	61	9.8	1274	4	US-09-949-016-8831	Sequence 8831, Ap
69	61	9.8	1274	4	US-09-949-016-8832	Sequence 8832, Ap
70	61	9.8	1749	4	US-09-640-419C-28	Sequence 28, App1
71	61	9.8	1839	2	US-09-172-977-4	Sequence 4, App1
72	61	9.8	1839	4	US-09-404-109-4	Sequence 4, App1
73	61	9.8	3924	4	US-09-538-092-1246	Sequence 1246, Ap
74	60.5	9.7	204	1	US-08-185-432-8	Sequence 8, App1
75	60.5	9.7	303	1	US-08-185-432-5	Sequence 5, App1
76	60.5	9.7	333	4	US-09-071-035-116	Sequence 116, App
77	60.5	9.7	406	4	US-09-071-035-114	Sequence 114, App
78	60.5	9.7	511	4	US-09-248-796A-14125	Sequence 14125, A
79	60.5	9.7	557	4	US-09-134-000C-4354	Sequence 4354, Ap
80	60.5	9.7	621	4	US-09-313-942-20	Sequence 20, App1
81	60.5	9.7	737	1	US-08-185-432-2	Sequence 2, App1
82	60.5	9.7	737	1	US-08-185-432-4	Sequence 4, App1
83	60.5	9.7	854	4	US-09-350-841A-1589	Sequence 1589, Ap
84	60.5	9.7	856	3	US-08-486-099-103	Sequence 103, App
85	60.5	9.7	856	3	US-08-484-223B-103	Sequence 103, App
86	60.5	9.7	856	3	US-08-919-597-103	Sequence 103, App
87	60.5	9.7	856	3	US-08-475-668A-103	Sequence 103, App
88	60.5	9.7	856	3	US-08-485-551A-103	Sequence 103, App
89	60.5	9.7	856	3	US-08-485-551A-103	Sequence 103, App
90	60.5	9.7	856	3	US-08-485-551A-103	Sequence 103, App
91	60.5	9.7	856	3	US-08-485-551A-103	Sequence 103, App
92	60.5	9.7	856	4	US-08-470-896-103	Sequence 103, App
93	60.5	9.7	856	4	US-08-485-546A-103	Sequence 103, App
94	60.5	9.7	856	4	US-08-487-266A-103	Sequence 103, App
95	60.5	9.7	857	1	US-08-320-151-10	Sequence 10, App1
96	60.5	9.7	857	1	US-08-413-118-10	Sequence 10, App1
97	60.5	9.7	857	3	US-08-804-439A-18	Sequence 18, App1
98	60.5	9.7	857	3	US-08-360-107A-113	Sequence 113, App
99	60.5	9.7	857	3	US-08-473-446-10	Sequence 10, App1
100	60.5	9.7	857	3	US-08-720-229-18	Sequence 18, App1

ALIGNMENTS

RESULT 1
US-09-334-220-2

```
; Sequence 2, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; FILE REFERENCE: 2427/0704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3461
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-334-220-2
```

```
Query Match 100.0%; Score 625; DB 3; Length 3461;
Best Local Similarity 100.0%; Pred. No. 1.5e-69;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EOCGITMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 60
DB 230 EOCGITMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 289
DB 290 NNTADWIOLEKIRAPNSVSTVHILYLPBEAKGESVQFQMKODSLVGEVYEACWAL 346
```

RESULT 2

```
US-09-334-220-1
; Sequence 1, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; FILE REFERENCE: 2427/0704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-334-220-1
```

```
Query Match 91.5%; Score 572; DB 3; Length 3460;
Best Local Similarity 90.6%; Pred. No. 8.8e-63;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 EOCGITMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 60
DB 229 EOCGITMGNNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK 288
QY 61 NNTADWIOLEKIRAPNSVSTVHILYLPBEAKGESVQFQMKODSLVGEVYEACWAL 117
DB 289 NNTADWIOLEKIRAPNSVSTVHILYLPBEAKGESVQFQMKODSLVGEVYEACWAL 345
```

RESULT 3
US-09-236-063-1

```
; Sequence 1, Application US/09236063
; Patent No. 6537792
; GENERAL INFORMATION:
; APPLICANT: Allen, Martin
; APPLICANT: Fang, Tsuei-Yun
; APPLICANT: Li, Yuxing
; APPLICANT: Liu, Hsuan-Liang
; APPLICANT: Chen, Hsui-Mei
; APPLICANT: Coutinho, Pedro
; APPLICANT: Hanzatko, Richard
; APPLICANT: Ford, Clark
; TITLE OF INVENTION: PROTEIN ENGINEERING OF GLUCOAMYLASE TO
; TITLE OF INVENTION: INCREASE PH OPTIMUM, SUBSTRATE SPECIFICITY AND
; TITLE OF INVENTION: THERMOSTABILITY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6537792thwestern Hwy.
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/236, 063
; FILING DATE:
; CLASSIFICATION:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0812.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus
US-09-236-063-1
```

```
Query Match 13.0%; Score 81; DB 4; Length 616;
Best Local Similarity 25.2%; Pred. No. 0.31;
Matches 34; Conservative 15; Mismatches 54; Indels 32; Gaps 4;
```

```
QY 9 GNAVTFCEPYGPRELTTTCLNTTASVLOFSIGSGSCRFPSYSDPSITVSXAK-- 61
DB 473 GGTITTAFTGSGSVSTSKTATASKTSTSTSTCTTPPAVAVTFDITATTGENTY 532
QY 62 -----NTADWIOLEKIRAPNSVSTVHILYLPBEAKGESVQF-----Q 99
DB 533 LVGSISQGLDWTSIGLMSADKYSTSPIMVVTYTLF---AGSPFYKFIRESDSVE 589
QY 100 WKQDSLRLVGEVYEAC 114
DB 590 WSDPNRREYVPOAC 604
```

RESULT 4
US-09-351-814-13

Sequence 13, Application US/09351814
Patent No. 6352851
GENERAL INFORMATION:
APPLICANT: Nielsen, Bjarne Roenfeldt
APPLICANT: Nielsen, Allan
APPLICANT: Pedersen, Henrik
APPLICANT: Vind, Jesper
APPLICANT: Hendriksen, Hanne Vang
APPLICANT: Frandsen, Torben Peter
TITLE OF INVENTION: Glucosylase Variants
FILE REFERENCE: 5636.200-US
CURRENT APPLICATION NUMBER: US/09/351,814
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: PA 1998 00937
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: PA 1998 01667
EARLIER FILING DATE: 1998-12-17
EARLIER APPLICATION NUMBER: 60/093,528
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/115,545
EARLIER FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PaetSEQ for Windows Version 3.0
SEQ ID NO 13
LENGTH: 640
TYPE: PRT
ORGANISM: ASPERGILLUS NIGER
US-09-351-814-13

Query Match 13.0%; Score 81; DB 3; Length 640;
Best Local Similarity 25.2%; Pred. No. 0.32;
Matches 34; Conservative 15; Mismatches 54; Indels 32; Gaps 4;

QY 9 GNAVTCEPYGRELTTTCNTTASVLPQISGSGCRFSYS-----DPSITVSYAKN-- 61
DB 497 GGTATTATPGSGSVSTSKTATYASKTSTSTSCCTPTAAVTFDLTATTYGENIY 556
QY 62 -----NTADWIQLEKIRAPSNVSTVIHILYPEBAKGSVQ-----Q 99
DB 557 LVGSISQLGDMETSDGIALSADKRTSSDPLMYTVLP---AGESEFYKIRIESDSV 613
QY 100 WKODSLRGEVVEAC 114
DB 614 WESDPNRBYTVPOAC 628

RESULT 5
US-09-248-796A-14918
Sequence 14918, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 14918
LENGTH: 568
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-14918

Query Match 11.3%; Score 70.5; DB 4; Length 568;
Best Local Similarity 31.5%; Pred. No. 5.9;
Matches 23; Conservative 16; Mismatches 15; Indels 19; Gaps 5;
QY 45 CRFSYSDPSITVSYAKNNT-----ADWIQ-LEKIRAPSNVSTVIHILYPEBAKGSVQ 97

DB 250 CKFTYRDDAIR-AYLA-NLTBEGLYKVDWTQNDKSNAPNVN-----EPQGESLJK 297
QY 98 PQWKODSLRGEV 110
DB 298 VKFDKXISFVGQL 310

RESULT 6
US-09-583-110-3587
Sequence 3587, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3587
LENGTH: 462
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3587

Query Match 10.9%; Score 68; DB 4; Length 462;
Best Local Similarity 24.6%; Pred. No. 9.1;
Matches 30; Conservative 16; Mismatches 42; Indels 34; Gaps 6;

QY 6 INHNAVTCEP-----YGPRELTTTCNTTAA-----SVLPQISGSGCRFSYS 50
DB 70 IFSKDVLFTPSISKLHLRLISYAKGEGTEIALGKDALMSKIMSGFGLSGFSQISRF 129
QY 51 DPSI---TYSYAKNNTADWIQLEKIRAPSNVSTVIH-----LYLPEBAKGSVQ 97
DB 130 VPSVLTRTVSQSFR-----MWSKV-VPOKEBDLHLMNPYQVLMLTPESEKAAAD 183
QY 98 PQ 99
DB 184 FE 185

RESULT 7
US-09-107-433-3273
Sequence 3273, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433

PRIOR FILING DATE: 30-Jun-1998
APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3273:
SEQUENCE CHARACTERISTICS:
LENGTH: 473 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..473
SEQUENCE DESCRIPTION: SEQ ID NO: 3273:
US-09-107-433-3273

Query Match 10.9%; Score 68; DB 4; Length 473;
Best Local Similarity 24.6%; Pred. No. 9.4;
Matches 30; Conservative 16; Mismatches 42; Indels 34; Gaps 6;

QY 6 IMHGNVTFCEP-----YGPRLTTCLNTTTA---SVLQFISGSGSCFSSYS 50
DB 81 IFSKRVLTSPISKLHLRLHSYAKGEGHIALGTXDMLGSKIMSGFSGRISHRF 140
QY 51 DESI---TTSYAKNNTADMIOLEKIRAPSNVSVIH-----LVPEAKGESVQ 97
DB 141 VSVLTRLTVSOSFNR-----MWSKV-VPOKEBDLHLMNOPYIVLMLTPESEKRAAD 194
QY 98 FQ 99
DB 195 FE 196
RESULT 8
US-09-902-540-13902
Sequence 13902, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 13902
LENGTH: 884
TYPE: PRF
ORGANISM: Myxococcus xanthus
US-09-902-540-13902

Query Match 10.9%; Score 68; DB 4; Length 884;
Best Local Similarity 25.0%; Pred. No. 24;
Matches 32; Conservative 15; Mismatches 51; Indels 30; Gaps 8;
QY 1 EOCGTIMHGNVTFCEP-----GPRELTTCLNT-----TTASVLFQFISGS 42

DB 603 EKVDAVYHNGALVNTFYEPESMRANVLGTRILALCVTRIKPLHYVSTSVL--PVGR 660
QY 43 GSGCFYSYDP-----STVSYAKNNTADMIOLEKIRAPSNVSVIHILYLPPEAKGESVQ 97
DB 661 -QLPREDPELEAASIVGVASQSK---WVAEKLVREASQREGI.PVTI--HRPGRVTGHSRT 715
QY 98 FQWKODSL 105
DB 716 GAMNTDDL 723

RESULT 9
US-09-341-461-2
Sequence 2, Application US/09341461
Patent No. 6586389
GENERAL INFORMATION:
APPLICANT: Hammond, Timothy G.
APPLICANT: Verrouet, Pierre J.
TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
FILE REFERENCE: D6148
CURRENT APPLICATION NUMBER: US/09/341,461
CURRENT FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: PCT/US99/01259
PRIOR FILING DATE: 1999-01-21
NUMBER OF SEQ ID NOS: 40
SEQ ID NO 2
LENGTH: 3623
TYPE: PRF
ORGANISM: rat
FEATURE:
OTHER INFORMATION: amino acid sequence of rat cubilin protein
US-09-341-461-2

Query Match 10.8%; Score 67.5; DB 4; Length 3623;
Best Local Similarity 27.5%; Pred. No. 2.3e+02;
Matches 38; Conservative 13; Mismatches 44; Indels 43; Gaps 10;

QY 3 CGTIMHGNVTFCEPYPREI--TTTC-----LNTTASVLFQFISG--SGSCR 46
DB 817 CGGMLRGEGF-FRSPFYNAVYGRTRCMTISQFORQVLLNFT-----DFQIGSSASCD 870
QY 47 FSYSD--PSITVSYAKNNTADMIOLEKI---RAPSNVSVIHILV-----PEAKGES 95
DB 871 TDYIEIGPSSVLSGPDN-----EKFCSSNIPSFITSVNVILVTFVKSSSMENRGPT 922
QY 96 VOQWKODSLRVGEVYEA 113
DB 923 AKF--SSDKLECGEVUTA 938

RESULT 10
US-09-802-371-2
Sequence 2, Application US/09802371
Patent No. 6723533
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Rudolph-Owen, Laura
TITLE OF INVENTION: 26934, A No. 6723533el Cytidine Deaminase-Like
FILE REFERENCE: 35800/213921
CURRENT APPLICATION NUMBER: US/09/802,371
CURRENT FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/188,294
PRIOR FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRF
ORGANISM: Homo sapiens
US-09-802-371-2

Query Match 10.7%; Score 67; DB 4; Length 339;
 Best Local Similarity 24.3%; Pred. No. 7.7;
 Matches 25; Conservative 21; Mismatches 39; Indels 18; Gaps 4;

QY 2 QCGTTHGNAVTCEPCEPRLTTTCNTTASVLOFSIGSGCRFSYSDPSITVSANKNTA---SDPSITV---56
 DB 112 QALIKRGSILKNCDDLFSRRKPCSA-----LKNIVAGVIRISYWPADPEISLITE 163

QY 57 -SYAKNNTADWIOLEKIRAPSNVSTVHIILYPEAKGESVQF 98
 DB 164 ASSEDEKADAKAVERLKNSRA---HVCVLQPLVCTMVQF 202

RESULT 11
 US-09-489-039A-11918
 ; Sequence 11918, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 11918
 ; LENGTH: 356
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-11918

Query Match 10.7%; Score 67; DB 4; Length 356;
 Best Local Similarity 26.3%; Pred. No. 8.3;
 Matches 31; Conservative 19; Mismatches 56; Indels 12; Gaps 4;

QY 9 GNAVTCEPCEPRLTTTCNTTASVLOFSIGSGCRFSYSDPSITVSANKNTA---DWI 67
 DB 238 GGAIAPKAGVQPLPPTVGNDAITIALQULIT-AGDQNTISKPSITVAALAAKAVAVDFI 296

QY 68 QLEKIRAPSNVSTVHIILYPEAKGESVQF-----WKDSLVEGVCEACMAL 117
 DB 297 QGKKPQASTLVNTPSQLFTPEVVTANIKAEIFDKKIQTW--DQVSGEYAAACQKL 352

RESULT 12
 US-09-514-245-22
 ; Sequence 22, Application US/09514245
 ; Patent No. 6703023
 ; GENERAL INFORMATION:
 ; APPLICANT: JESTIN, Andre
 ; APPLICANT: ALBINA, Emmanuel
 ; APPLICANT: Le CANN, Pierre
 ; APPLICANT: BLANCHARD, Philippe
 ; APPLICANT: HUTTET, Beelyne
 ; APPLICANT: ARNAUD, Claire
 ; APPLICANT: TRUONG, Catherine
 ; APPLICANT: MAHE, Dominique
 ; APPLICANT: CARIOLET, Roland
 ; APPLICANT: MADEC, Francois
 ; TITLE OF INVENTION: CIRCOVIRUS SEQUENCES ASSOCIATED WITH PIGLET WEIGHT LOSS DISEASE
 ; FILE REFERENCE: 065691/016
 ; CURRENT APPLICATION NUMBER: US/09/514, 245
 ; CURRENT FILING DATE: 2000-02-28
 ; PRIOR APPLICATION NUMBER: FR 97/15396
 ; PRIOR FILING DATE: 1997-12-05
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Type B PWD circovirus

US-09-514-245-22
 Query Match 10.6%; Score 66.5; DB 4; Length 569;
 Best Local Similarity 24.4%; Pred. No. 19;
 Matches 32; Conservative 14; Mismatches 34; Indels 51; Gaps 8;

QY 5 TINEGNAVTCEPCEPRLTTTCNTTASVLOFSIGSGCRFSYSDPSITVSANKNTA---SDPSITV---61
 DB 222 TTFHNNNIVKIAQGP-----IIQSQTSIWQ-----SYL--SFTSSYKQGAT 263

QY 62 -----NTADWIOLEKIRAPSN-----VSTVHIILYPEAKGESVQF 100
 DB 264 NONGAILGROFPVGSQSDMSYFSKI--PPNSGYKPLISCFGLFPALEDGKGGWAFKX 321

QY 101 -----KQDS 104
 DB 322 IFWIVSDKQDS 332

RESULT 13
 US-08-409-995-5
 ; Sequence 5, Application US/08409995
 ; Patent No. 5646259
 ; GENERAL INFORMATION:
 ; APPLICANT: Barenkamp, Stephen I.
 ; APPLICANT: St. Geme III, Joseph W.
 ; TITLE OF INVENTION: Haemophilus Adhesion Proteins
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 ; STREET: Four Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/409, 995
 ; FILING DATE: 24-MAR-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Silva, Robin M.
 ; REGISTRATION NUMBER: 38,304
 ; REFERENCE/DOCKET NUMBER: A-61053/RFT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 658 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 US-08-409-995-5

Query Match 10.6%; Score 66; DB 1; Length 658;
 Best Local Similarity 24.2%; Pred. No. 28;
 Matches 24; Conservative 19; Mismatches 36; Indels 20; Gaps 4;

QY 1 ECGTTHGNAVTCEPCEPRLTTTCNTTASVLOFSIGSGCRFSYSDPSITVSANKNTA---DWI 67
 DB 129 EKSQGVYHAEVLF-BGKGGVQVTSSENGKHTITPALAKDLGVKATVADTLITIGGAA 187

QY 46 RFSYSDPSITVSANKNTADWIOLEKIRAPSNVSTVHI 84
 DB 188 AGATTPKAVVT-----STTDGLKFAKDAGANGDTIVHL 222

```

RESULT 14
US-08-685-467-5
: Sequence 5, Application US/08685467
: Patent No. 6060059
: GENERAL INFORMATION:
: APPLICANT: St. Gene III, Joseph W.
: APPLICANT: Barenkamp, Stephen J.
: TITLE OF INVENTION: HAEMOPHIUS ADHESION PROTEINS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/685,467
: FILING DATE: 22-JUL-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/409,995
: FILING DATE: 24-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Silva, Robin M.
: REGISTRATION NUMBER: 38,304
: REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 658 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
:
US-08-685-467-5
:
Query Match 10.6%; Score 66; DB 3; Length 658;
Best Local Similarity 24.2%; Pred. NO. 28;
Matches 24; Conservative 19; Mismatches 36; Indels 20; Gaps 4
:
CY 1 ECGGTGAGCAVTFCEPYGPRELITTC-----LNTTASVLO-FSIGSGSC 45
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 EKSQOVKRADEVLV-FEGKGQVQVSTSENKHTIFALAKDQGVKATVSDTLITGGAA 187
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
CY 46 RFSYSDPSITVYAKNNTADWITQLEKIRAPSNVSTVIHI 84
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 AGATTTPKRVNT---STTDGLKFAKDAAGANGDTTVHL 222
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-08-913-942-5
: Sequence 5, Application US/08913942
: Patent No. 6200578
: GENERAL INFORMATION:
: APPLICANT: St. Gene, Joseph
: APPLICANT: Barenkamp, Stephen J.
: TITLE OF INVENTION: HAEMOPHIUS ADHESION PROTEINS
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States

```

```

ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,942
FILING DATE: 29-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/4031
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vance, Dolly A.
REGISTRATION NUMBER: 39,054
REFERENCE//DOCKET NUMBER: A-61053-1/RFT/RMS/DAV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-913-942-5

Query Match          10.6%; Score 66; DB 3; Length 658;
Best Local Similarity 24.2%; Pred. No. 28;
Matches      24; Conservative    19; Mismatches     36; Indels   20; Gaps    4;

QY      1 EOCGTINMGNAVTPCEPYGPRLTTTC-----LNTTASVLQ-FSIGSGSC 45
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      129 EKSGQVGHADEVLFE-EKGKGVQTSTSENGKHITPALAKDLGKVATVSDTLRTGGGA 187

QY      46 RFSYSDPSTIVSYAKNNTADMIQLKRAISNVSVTHI 84
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      188 AGATTPKVNVT---STTDGLKPAKDAAAGNDPTVHL 222

Search completed: August  4, 2005, 15:37:01
Job time : 45 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 15:35:34 ; Search time 158 Seconds
(without alignments)
289.092 Million cell updates/sec

Title: US-09-897-438b-2

Perfect score: 625

Sequence: 1 ECGGTIMGNATVFCPEYGP.....FQWKQSLRVEYEAQWAL 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1752860 seqs, 390397842 residues

Total number of hits satisfying chosen parameters: 1752860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	100.0	117	9	US-09-897-438b-2
2	625	100.0	117	9	US-09-832-189a-4
3	572	91.5	3427	16	US-10-408-765A-1624
4	572	91.5	3460	10	US-10-723-860-3739
5	572	91.5	3460	17	US-10-753-267-46
6	572	91.5	3460	20	US-11-019-829-22
7	499	79.8	432	10	US-09-832-189a-2
8	224	35.8	84	14	US-10-029-386-3034
9	190	30.4	70	9	US-09-864-761-47165
10	107	17.1	28	9	US-09-864-761-44551
11	83	13.3	316	16	US-10-767-701-32303

12	81	13.0	640	14	US-10-038-723-13	Sequence 13, Appl
13	81	13.0	640	15	US-10-421-586-14	Sequence 14, Appl
14	74	11.8	615	18	US-10-991-654-6	Sequence 6, Appl
15	74	11.8	639	18	US-10-991-654-5	Sequence 5, Appl
16	73	11.7	139	15	US-10-424-599-212843	Sequence 212843
17	73	11.7	1095	14	US-10-228-063-45	Sequence 45, Appl
18	72	11.5	208	15	US-10-425-114-71644	Sequence 71644, A
19	72	11.5	247	16	US-10-425-115-229189	Sequence 229189, A
20	71.5	11.4	91	14	US-10-029-386-29362	Sequence 29362, A
21	71	11.4	1051	16	US-10-425-115-338413	Sequence 338413, A
22	70.5	11.3	577	15	US-10-425-114-43051	Sequence 43051, A
23	70.5	11.3	901	15	US-10-424-599-193645	Sequence 193645, A
24	70	11.2	222	17	US-10-857-673-12	Sequence 12, Appl
25	70	11.2	324	17	US-10-857-673-4	Sequence 4, Appl
26	70	11.2	366	15	US-10-425-114-56131	Sequence 56131, A
27	70	11.2	593	15	US-10-424-599-222406	Sequence 222406
28	70	11.2	809	17	US-10-857-673-2	Sequence 2, Appl
29	70	11.2	809	17	US-10-857-673-15	Sequence 15, Appl
30	69	11.2	175	16	US-10-425-115-230792	Sequence 230792, A
31	68.5	11.0	326	15	US-10-320-797-3095	Sequence 3095, Ap
32	68.5	11.0	572	15	US-10-389-566-2229	Sequence 2229, Ap
33	68.5	11.0	611	15	US-10-389-566-1060	Sequence 1060, Ap
34	68.5	11.0	669	15	US-10-424-599-115992	Sequence 115992, A
35	68.5	11.0	1124	16	US-10-437-963-159628	Sequence 159628, A
36	68	10.9	426	14	US-10-214-524-27	Sequence 27, Appl
37	68	10.9	472	17	US-10-472-928-4158	Sequence 4158, Ap
38	68	10.9	473	18	US-10-617-320-3373	Sequence 3273, Ap
39	68	10.9	513	16	US-10-739-930-6657	Sequence 6657, Ap
40	68	10.9	554	17	US-10-850-270-8	Sequence 8, Appl
41	68	10.9	573	15	US-10-104-047-3017	Sequence 3017, Ap
42	68	10.9	791	16	US-10-437-963-109617	Sequence 109617, A
43	68	10.9	932	16	US-10-437-963-190783	Sequence 190783, A
44	67.5	10.8	1165	16	US-10-425-115-359139	Sequence 359139, A
45	67	10.7	222	11	US-09-833-245-1776	Sequence 1776, Ap
46	67	10.7	267	11	US-09-833-245-1779	Sequence 1779, Ap
47	67	10.7	339	9	US-09-802-371-2	Sequence 2, Appl
48	67	10.7	339	16	US-10-785-135-2	Sequence 2, Appl
49	67	10.7	489	11	US-09-833-245-1778	Sequence 1778, Ap
50	67	10.7	633	16	US-10-437-963-11041	Sequence 141041, A
51	67	10.7	743	15	US-10-418-836-35	Sequence 35, Appl
52	67	10.7	743	20	US-11-007-886-35	Sequence 35, Appl
53	66.5	10.6	544	15	US-10-282-122A-53715	Sequence 53715, A
54	66.5	10.6	569	15	US-10-682-420-22	Sequence 22, Appl
55	66.5	10.6	569	15	US-10-409-613-22	Sequence 22, Appl
56	66.5	10.6	569	15	US-10-442-180-22	Sequence 22, Appl
57	66.5	10.6	569	17	US-10-718-266-22	Sequence 22, Appl
58	66.5	10.6	569	17	US-10-775-337-22	Sequence 22, Appl
59	66.5	10.6	569	17	US-10-637-011-22	Sequence 22, Appl
60	66	10.6	128	16	US-10-425-115-303289	Sequence 303289, A
61	66	10.6	172	16	US-10-767-701-45933	Sequence 45933, A
62	66	10.6	223	17	US-10-857-673-9	Sequence 9, Appl
63	66	10.6	367	16	US-10-425-115-256502	Sequence 256502, A
64	66	10.6	743	15	US-10-424-599-222404	Sequence 222404, A
65	66	10.6	743	20	US-10-418-836-31	Sequence 31, Appl
66	66	10.6	743	20	US-11-007-886-31	Sequence 31, Appl
67	66	10.6	818	17	US-10-857-673-10	Sequence 10, Appl
68	66	10.6	971	17	US-10-732-923-13599	Sequence 13599, A
69	66	10.6	1098	9	US-09-797-862-32	Sequence 32, Appl
70	66	10.6	1098	17	US-10-637-659-32	Sequence 32, Appl
71	65.5	10.5	139	16	US-10-425-115-276953	Sequence 276953, A
72	65.5	10.5	160	15	US-10-424-599-207598	Sequence 207598, A
73	65.5	10.5	220	16	US-10-425-115-207572	Sequence 207572, A
74	65.5	10.5	312	15	US-10-369-493-12839	Sequence 12839, A
75	65.5	10.5	465	14	US-10-128-71A-3038	Sequence 3038, Ap
76	65.5	10.5	549	9	US-09-738-626-4536	Sequence 4536, Ap
77	65.5	10.5	549	13	US-10-051-902-30	Sequence 30, Appl
78	65.5	10.5	549	13	US-10-051-909-30	Sequence 30, Appl
79	65.5	10.5	549	13	US-10-389-566-2077	Sequence 2077, Ap
80	65.5	10.5	611	15	US-10-187-975-110	Sequence 110, App
81	65.5	10.5	620	15	US-10-128-71A-8038	Sequence 8038, Ap
82	65.5	10.5	624	14	US-10-739-930-6228	Sequence 6228, Ap
83	65.5	10.5	953	16	US-10-210-130-100	Sequence 100, App
84	65.5	10.5	1738	15	US-10-210-130-100	Sequence 100, App

85 65 10.4 223 17 US-10-857-673-18 Sequence 18, Appl
86 65 10.4 295 16 US-10-425-115-233981 Sequence 233981,
87 65 10.4 438 9 US-09-738-626-6703 Sequence 6703, Ap
88 65 10.4 451 14 US-09-712-363-171 Sequence 171, App
89 65 10.4 451 14 US-10-080-170-400 Sequence 400, App
90 65 10.4 451 14 US-10-080-170-400 Sequence 400, App
91 65 10.4 451 16 US-10-468-356-400 Sequence 400, App
92 65 10.4 4823 15 US-10-051-874-159 Sequence 163, App
93 64.5 10.3 109 16 US-10-437-963-118926 Sequence 118926,
94 64.5 10.3 389 15 US-10-425-114-54288 Sequence 54288, A
95 64.5 10.3 550 9 US-09-815-242-10499 Sequence 10499, A
96 64.5 10.3 559 16 US-10-425-115-321798 Sequence 321798,
97 64.5 10.3 599 14 US-10-138-075-5 Sequence 5, Appli
98 64.5 10.3 599 16 US-10-739-930-10739 Sequence 10739, A
99 64.5 10.3 507 16 US-10-437-963-196952 Sequence 196952,
100 64.5 10.3 1001 16 US-10-437-963-196954 Sequence 196954,

ALIGNMENTS

RESULT 1
US-09-897-438b-2
; Sequence 2, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPITOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-897-438b-2
Query Match 100.0%; Score 625; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 3,4e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 1 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
QY 61 NNTADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 117
DB 61 NNTADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 117
RESULT 2
US-09-832-189A-4
; Sequence 4, Application US/09832189A
; Publication No. US20030114657A1
; GENERAL INFORMATION:
; APPLICANT: RIKEN
; TITLE OF INVENTION: Truncated Reelin Protein and DNA Encoding the Same
; FILE REFERENCE: PH-1167
; CURRENT APPLICATION NUMBER: US/09/832,189A
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: JP 2000-109954
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mus musculus

US-09-832-189A-4
Query Match 100.0%; Score 625; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 2,8e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 230 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 289
QY 61 NNTADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 117
DB 290 NNTADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 346

RESULT 3
US-10-408-765A-1624
; Sequence 1624, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1624
; LENGTH: 3427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124,
; LOCATION: 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
; LOCATION: 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148,
; LOCATION: 149, 150, 151, 152, 153, 154, 155, 156, 157, 158
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-1624
Query Match 91.5%; Score 572; DB 16; Length 3427;
Best Local Similarity 90.6%; Pred. No. 3,9e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 229 EOCGTIMGNNAVTFCEPYGPRELTTTCNTTASVLOFSIGSGCRFSYSDPSITVSYAK 288
QY 61 NNTADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 117
DB 289 NNSADWIOLEKIRAPSNVSTVIHIIYLPEBAKGSVQFQWKDLSRVGEVYEACMAL 345
RESULT 4
US-10-723-860-3739
; Sequence 3739, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPIS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26

```

; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3739
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-3739

Query Match          91.5%; Score 572; DB 16; Length 3460;
Best Local Similarity 90.6%; Pred. No. 3.9e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVSYAK 60
DB 229 EOCGAIMHGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVYAK 288

QY 61 NNTADWIOLEKIRAPSNVSTVTHIILYLPBEANGESVQFQWKDLSRVEVYEAQWAL 117
DB 289 NNSADWIOLEKIRAPSNVSTIHIILYLPEDAGENVQFQWKDNLKRVGEVYEAQWAL 345

RESULT 5
US-10-753-267-46
; Sequence 46, Application US/10753267
; Publication No. US20050037946A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Steagliano, Nancy E.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Donoghue, Mary A.
; APPLICANT: Rodrique-Way, Amelie
; APPLICANT: Tomlinson, James E.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,
; TITLE OF INVENTION: 10653, 9255, 21668, 17794, 2210, 6169, 10102, 21061, 17662,
; TITLE OF INVENTION: 1468, 12822, 6350, 9035, 1820, 23652, 7301, 8925, 8701,
; TITLE OF INVENTION: 3533, 9462, 9123, 12788, 17729, 65552, 1261, 21476, 33770,
; TITLE OF INVENTION: 9380, 2569654, 33556, 53656, 44143, 32612, 10671, 261,
; TITLE OF INVENTION: 44570, 41922, 2552, 2417, 19319, 43969, 8921, 8993, 955,
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419
; FILE REFERENCE: MP103-003P1RNOMNIM
; CURRENT APPLICATION NUMBER: US/10753,267
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 60/439,683
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/445,216
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,036
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/454,189
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/457,541
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/466,411
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/469,041
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/477,414
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3460
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-753-267-46

Query Match          91.5%; Score 572; DB 17; Length 3460;
Best Local Similarity 90.6%; Pred. No. 3.9e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVSYAK 60
DB 229 EOCGAIMHGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVYAK 288

QY 61 NNTADWIOLEKIRAPSNVSTVTHIILYLPBEANGESVQFQWKDLSRVEVYEAQWAL 117
DB 289 NNSADWIOLEKIRAPSNVSTIHIILYLPEDAGENVQFQWKDNLKRVGEVYEAQWAL 345

RESULT 6
US-11-019-829-22
; Sequence 22, Application US/11019829
; Publication No. US20050136465A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Novel targets for obesity from subcutaneous fat
; FILE REFERENCE: 22304
; CURRENT APPLICATION NUMBER: US/11/019,829
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: reelin (RELN)
; LOCATION: (1)..(3460)
; OTHER INFORMATION: U79716
US-11-019-829-22

Query Match          91.5%; Score 572; DB 20; Length 3460;
Best Local Similarity 90.6%; Pred. No. 3.9e-56;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVSYAK 60
DB 229 EOCGAIMHGNNAVTFCEPYGPRELTTTCLNTTTASVLOFSIGSGSCRFSSDPSITVYAK 288

QY 61 NNTADWIOLEKIRAPSNVSTVTHIILYLPBEANGESVQFQWKDLSRVEVYEAQWAL 117
DB 289 NNSADWIOLEKIRAPSNVSTIHIILYLPEDAGENVQFQWKDNLKRVGEVYEAQWAL 345

RESULT 7
US-09-832-189A-2
; Sequence 2, Application US/09832189A
; Publication No. US20030114657A1
; GENERAL INFORMATION:
; APPLICANT: RIKEN
; TITLE OF INVENTION: Truncated Reelin Protein and DNA Encoding the Same
; FILE REFERENCE: PH-1167
; CURRENT APPLICATION NUMBER: US/09/832,189A
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: JP 2000-109954
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-832-189A-2

Query Match          79.8%; Score 499; DB 10; Length 432;
```

```
Best Local Similarity 74.4%; Pred. No. 7.5e+49;
Matches 87; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

QY 1 EOCGIMGNNAVTCBPYGPRLTTTCLNTTASVLOFSIGSSCRFSYSDBSITVSYAK 60
Db 231 EOCGAIMHGAVTFCBPYGPRLITVQNMTTASVLOFSIGSSCRFSYSDBSITVSYAK 290

QY 61 NNTAAMIOLEKTRAPSNVSTVTHILYLPPEAKGESVQFQWKDSLAVGEYEAACMAL 117
Db 291 NNSSSMPLERISAPSNVSTIITHILYLPPEAKGENVCFRWMQADGVYEACMAL 347

RESULT 8
US-10-029-386-30334
; Sequence 30334, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30334
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR7.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: SWISSPROT HIT: O14063, EVALUE 2.00e+00
US-10-029-386-30334

Query Match 35.8%; Score 224; DB 14; Length 84;
Best Local Similarity 88.9%; Pred. No. 6.7e-18;
Matches 40; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 73 RAPSNTVTHILYLPPEAKGESVQFQWKDSLAVGEYEAACMAL 117
Db 3 RAPSNTIITHILYLPPEAKGENVCFRWMQADGVYEACMAL 47

RESULT 9
US-09-864-761-47165
; Sequence 47165, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
```

```
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47165
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
; OTHER INFORMATION: SWISSPROT HIT: O14063, EVALUE 1.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE938667.1, EVALUE 2.00e-22
US-09-864-761-47165

Query Match 30.4%; Score 190; DB 9; Length 70;
Best Local Similarity 86.8%; Pred. No. 4.6e-14;
Matches 33; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 80 TVIHTILYLPPEAKGESVQFQWKDSLAVGEYEAACMAL 117
Db 1 TIIHTILYLPPEAKGENVCFRWMQADGVYEYEAACMAL 38

RESULT 10
US-09-864-761-44551
; Sequence 44551, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
```

PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 44551
LENGTH: 28
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC000121.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.44
OTHER INFORMATION: SWISSPROT HIT: P76008, EVALU 2.90e+00
US-09-864-761-44551

Query Match 17.1%; Score 107; DB 9; Length 28;
Best Local Similarity 94.7%; Pred. No. 5.9e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOCGTIMGNATVFCPEYG 19
DB 10 EOCGAIMGNATVFCPEYG 28

RESULT 11
US-10-767-701-32303
Sequence 32303, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 32303
LENGTH: 316
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(316)

OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C124217_1.pcp
US-10-767-701-32303

Query Match 13.3%; Score 83; DB 16; Length 316;
Best Local Similarity 25.2%; Pred. No. 0.84;
Matches 34; Conservative 15; Mismatches 54; Indels 32; Gaps 4;

QY 9 GNAVTCBPYGRRLTTCNTTASVLOPSIGSGCRFSYS-----DSITVSYAKN-- 61
DB 172 GGTITATPTGSGSVSTSKTATASKTSTSSASCTTPTAVAVFDLTATTYGENY 231

QY 62 -----NTADWIOLEKIRAPSNVSTVHILVLPBEAKGESVOF-----Q 99
DB 232 LVGSISQLDQWETSDDIALSADKYTSSDPLMVTYVLP---AGESEYKFIRESDDSV 288

QY 100 WKODSLRVEGYEAC 114
DB 289 WESDPNREYTVPOAC 303

RESULT 12
US-10-038-723-13
Sequence 13, Application US/10038723
Publication No. US20030032163A1
GENERAL INFORMATION:
APPLICANT: Nielsen, Bjarne Roenfeldt
APPLICANT: Svendsen, Allan
APPLICANT: Pedersen, Henrik
APPLICANT: Vind, Jesper
APPLICANT: Hendriksen, Hanne Vang
APPLICANT: Frandsen, Torben Peter
TITLE OF INVENTION: Glucosylase Variants
FILE REFERENCE: 5636.200-US
CURRENT FILING DATE: 2002-01-02
PRIOR APPLICATION NUMBER: US/10/038,723
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/351,814
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: PA 1998 00937
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 01667
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,528
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/115,545
PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 640
TYPE: PRT
ORGANISM: ASPERGILLUS NIGER
US-10-038-723-13

Query Match 13.0%; Score 81; DB 14; Length 640;
Best Local Similarity 25.2%; Pred. No. 3.6;
Matches 34; Conservative 15; Mismatches 54; Indels 32; Gaps 4;

QY 9 GNAVTCBPYGRRLTTCNTTASVLOPSIGSGCRFSYS-----DSITVSYAKN-- 61
DB 497 GGTITATPTGSGSVSTSKTATASKTSTSSASCTTPTAVAVFDLTATTYGENY 556

QY 62 -----NTADWIOLEKIRAPSNVSTVHILVLPBEAKGESVOF-----Q 99
DB 557 LVGSISQLDQWETSDDIALSADKYTSSDPLMVTYVLP---AGESEYKFIRESDDSV 613

QY 100 WKODSLRVEGYEAC 114
DB 614 WESDPNREYTVPOAC 628

RESULT 13

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 15:29:54 ; Search time 492 Seconds
(without alignments)
277.758 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 625

Sequence: 1 ECGRTIMGNATVFCPEYGP.....FGWKQSLRGEVYECMAL 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/prodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/prodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/prodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/prodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/prodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/prodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/prodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/prodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/prodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/prodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/prodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/prodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/prodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/prodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/prodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/prodata/1/paa/US097B_COMB.pep:*
23: /cgn2_6/prodata/1/paa/US098_COMB.pep:*
24: /cgn2_6/prodata/1/paa/US099A_COMB.pep:*
25: /cgn2_6/prodata/1/paa/US099B_COMB.pep:*
26: /cgn2_6/prodata/1/paa/US100_COMB.pep:*
27: /cgn2_6/prodata/1/paa/US101_COMB.pep:*
28: /cgn2_6/prodata/1/paa/US102_COMB.pep:*
29: /cgn2_6/prodata/1/paa/US103_COMB.pep:*
30: /cgn2_6/prodata/1/paa/US104_COMB.pep:*
31: /cgn2_6/prodata/1/paa/US105_COMB.pep:*
32: /cgn2_6/prodata/1/paa/US106_COMB.pep:*
33: /cgn2_6/prodata/1/paa/US107_COMB.pep:*
34: /cgn2_6/prodata/1/paa/US108_COMB.pep:*
35: /cgn2_6/prodata/1/paa/US109_COMB.pep:*
36: /cgn2_6/prodata/1/paa/US110_COMB.pep:*
37: /cgn2_6/prodata/1/paa/US160_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB	ID	Description
1	625	100.0	117	23	US-09-897-438B-2	Sequence 2, Appli
2	625	100.0	589	23	US-09-832-189A-4	Sequence 4, Appli
3	625	100.0	3461	28	US-10-219-051B-6041	Sequence 6041, Ap
4	572	91.5	2119	21	US-09-724-676A-78129	Sequence 78129, A
5	572	91.5	2119	21	US-09-724-676A-78129	Sequence 78129, A
6	572	91.5	2830	37	US-60-245-221-81	Sequence 81, Appl
7	572	91.5	2830	37	US-60-258-279-28	Sequence 28, Appl
8	572	91.5	2954	24	US-09-949-003C-5671	Sequence 3671, Ap
9	572	91.5	2992	27	US-10-170-205E-37739	Sequence 37739, A
10	572	91.5	2992	37	US-60-453-050-10999	Sequence 10999, A
11	572	91.5	2992	37	US-60-453-135-10999	Sequence 10999, A
12	572	91.5	2992	37	US-60-466-412-10999	Sequence 10999, A
13	572	91.5	3107	21	US-09-724-676-78128	Sequence 78128, A
14	572	91.5	3107	21	US-09-724-676-78128	Sequence 78128, A
15	572	91.5	3427	30	US-10-408-765A-1624	Sequence 1624, Ap
16	572	91.5	3427	30	US-10-408-765A-1624	Sequence 1624, Ap
17	572	91.5	3427	37	US-60-389-987-1624	Sequence 1624, Ap
18	572	91.5	3427	37	US-60-412-418-1624	Sequence 1624, Ap
19	572	91.5	3460	1	PCT-US03-38193-3739	Sequence 3739, Ap
20	572	91.5	3460	1	PCT-US04-00393-46	Sequence 2394, Ap
21	572	91.5	3460	24	US-09-949-003C-2384	Sequence 6043, Ap
22	572	91.5	3460	28	US-10-219-051B-6043	Sequence 104, App
23	572	91.5	3460	31	US-10-505-928-104	Sequence 3739, Ap
24	572	91.5	3460	33	US-10-723-860-3739	Sequence 46, Appl
25	572	91.5	3460	33	US-10-753-267-46	Sequence 22, Appl
26	572	91.5	3460	36	US-11-019-829-22	Sequence 55656, A
27	572	91.5	3470	1	PCT-US01-08631-55656	Sequence 2, Appli
28	499	79.8	432	23	US-09-832-189A-2	Sequence 30334, A
29	224	35.8	84	26	US-10-029-386-30334	Sequence 47165, A
30	190	30.4	70	23	US-09-864-761-47165	Sequence 33869, A
31	190	30.4	70	23	US-10-203-135-33865	Sequence 78132, A
32	109.5	17.5	1504	21	US-09-724-676-78132	Sequence 78132, A
33	109.5	17.5	1504	21	US-09-724-676A-78132	Sequence 78131, A
34	109.5	17.5	2492	21	US-09-724-676-78131	Sequence 78131, A
35	109.5	17.5	2492	21	US-09-724-676A-78131	Sequence 78130, A
36	109.5	17.5	2863	21	US-09-724-676-78130	Sequence 78130, A
37	109.5	17.5	2863	21	US-09-724-676A-78130	Sequence 78130, A
38	107	17.1	28	23	US-09-864-761-44551	Sequence 44551, A
39	107	17.1	28	27	US-10-182-993-33489	Sequence 33489, A
40	107	17.1	28	28	US-10-203-134-33447	Sequence 33447, A
41	107	17.1	28	28	US-10-203-135-33725	Sequence 33725, A
42	107	17.1	28	28	US-10-203-136-33566	Sequence 33566, A
43	83	13.3	316	38	US-10-767-701-32303	Sequence 32303, A
44	82.5	13.2	1027	37	US-60-230-435-1085	Sequence 1085, Ap
45	82.5	13.2	1169	37	US-60-212-656-449	Sequence 449, App
46	81.5	13.0	1789	37	PCT-US01-08631-55654	Sequence 55654, A
47	81.5	13.0	1789	37	US-60-207-315-370	Sequence 370, App
48	81.5	13.0	1789	37	US-60-230-435-1413	Sequence 1413, Ap
49	81	13.0	614	22	US-09-791-537-78148	Sequence 78148, A
50	81	13.0	616	1	PCT-US00-00532-1	Sequence 1, Appli
51	81	13.0	640	1	PCT-US04-13841-4	Sequence 4, Appli
52	81	13.0	640	20	US-09-612-489-13	Sequence 13, Appl
53	81	13.0	640	22	US-09-791-537-113621	Sequence 113621, A
54	81	13.0	640	26	US-10-038-723-13	Sequence 13, Appl
55	81	13.0	640	30	US-10-421-58E-14	Sequence 118, App
56	77	12.3	574	1	PCT-US04-10222-118	Sequence 118, App
57	75	12.0	86	33	US-10-767-795-81672	Sequence 81672, A
58	74.5	11.9	834	1	PCT-US01-08631-55655	Sequence 55655, A
59	74	11.8	639	22	US-09-791-537-113620	Sequence 113620, A
60	74	11.8	640	22	US-09-791-537-152797	Sequence 152797, A
61	73	11.7	139	30	US-10-424-599-212843	Sequence 212843, A
62	73	11.7	639	22	US-09-791-537-113623	Sequence 113623, A
63	73	11.7	1095	1	PCT-US02-27122-45	Sequence 45, Appl
64	73	11.7	1095	1	PCT-US04-07182A-45	Sequence 45, Appl
65	73	11.7	1095	28	US-10-228-063-45	Sequence 45, Appl
66	72	11.5	208	30	US-10-219-999-59707	Sequence 59707, A
67	72	11.5	208	38	US-10-425-114-71644	Sequence 71644, A
68	72	11.5	208	30	US-10-425-114A-71644	Sequence 71644, A
69	72	11.5	208	37	US-60-324-109-25684	Sequence 25684, A

```
70 72 11.5 247 30 US-10-425-115-229189 Sequence 229189,
71 72 11.5 285 21 US-09-708-427-83671 Sequence 83672, A
72 72 11.5 350 21 US-09-708-427-83671 Sequence 83671, A
73 72 11.4 405 21 US-09-708-427-83670 Sequence 83670, A
74 71.5 11.5 91 26 US-10-029-386-29362 Sequence 29362, A
75 71 11.4 197 26 US-10-078-725-169 Sequence 169, App
76 71 11.4 197 37 US-60-270-153-169 Sequence 169, App
77 71 11.4 455 1 PCT-US99-22855-2315 Sequence 2315, Ap
78 71 11.4 457 1 PCT-US99-22855-2314 Sequence 2314, Ap
79 71 11.4 482 1 PCT-US99-22855-2313 Sequence 2313, Ap
80 71 11.4 1201 30 US-10-425-115-338413 Sequence 14321, A
81 71 11.4 1201 30 US-60-638-099-14321 Sequence 9116, Ap
82 70.5 11.3 271 30 US-10-415-182A-9116 Sequence 9116, Ap
83 70.5 11.3 279 30 US-10-415-182A-1146 Sequence 8170, Ap
84 70.5 11.3 400 27 US-10-179-131-8170 Sequence 8170, Ap
85 70.5 11.3 568 32 US-10-603-113-14918 Sequence 14918, A
86 70.5 11.3 568 37 US-60-096-409-14918 Sequence 14918, A
87 70.5 11.3 577 28 US-10-219-999-36487 Sequence 36487, A
88 70.5 11.3 577 30 US-10-425-114-43051 Sequence 43051, A
89 70.5 11.3 577 30 US-10-425-114A-43051 Sequence 43051, A
90 70.5 11.3 577 37 US-10-424-599-193645 Sequence 22439, A
91 70.5 11.3 901 30 US-10-424-599-193645 Sequence 193645,
92 70 11.2 222 34 US-10-857-673-12 Sequence 12, Appl
93 70 11.2 224 34 US-10-857-673-4 Sequence 4, Appl
94 70 11.2 366 28 US-10-219-999-47541 Sequence 47541, A
95 70 11.2 366 30 US-10-425-114-56131 Sequence 56131, A
96 70 11.2 366 30 US-10-425-114A-56131 Sequence 56131, A
97 70 11.2 366 37 US-60-324-109-30959 Sequence 30959, A
98 70 11.2 593 30 US-10-424-599-222406 Sequence 222406,
99 70 11.2 809 34 US-10-857-673-2 Sequence 2, Appl
100 70 11.2 809 34 US-10-857-673-15 Sequence 15, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-897-438b-2
; Sequence 2, Application US/09897438B
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPI TOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-897-438b-2
```

```
Query Match 100.0%; Score 625; DB 23; Length 117;
Best Local Similarity 100.0%; Pred. No. 2e-66;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 1 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 117
DB 61 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 117
```

```
RESULT 2
US-09-832-189A-4
; Sequence 4, Application US/09832189A
; GENERAL INFORMATION:
```

```
; APPLICANT: RIKEN
; TITLE OF INVENTION: Truncated Reelin Protein and DNA Encoding the Same
; FILE REFERENCE: PH-1167
; CURRENT APPLICATION NUMBER: US/09/832,189A
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: JP 2000-109954
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-832-189A-4

Query Match 100.0%; Score 625; DB 23; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.8e-65;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 230 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 289
QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 117
DB 290 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 346
```

```
RESULT 3
US-10-219-051B-6041
; Sequence 6041, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 6041
; LENGTH: 3461
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Refseq / NP_035391
; DATABASE ENTRY DATE: 2002-09-19
US-10-219-051B-6041
```

```
Query Match 100.0%; Score 625; DB 28; Length 3461;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 60
DB 230 EOCGTTMGNANTFCBPYGPRLTTTCNTTTTASVLOFSIGSGCRFSYSDPSITVSYAK 289
QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 117
DB 290 NNTADMIOLEKIRAPSNVSTVHIILYLPEBAKGSVQFQKODSLRVGEVYACMAL 346
```

```
RESULT 4
US-09-724-676-78129
; Sequence 78129, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
```



```

? TITLE OPINION: Variants of alternative splicing
? FILE REFERENCE: 129181.4 Compugen
? CURRENT APPLICATION NUMBER: US/09/724,676
? CURRENT FILING DATE: 2000-11-28
? NUMBER OF SEQ ID NOS: 97222
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 78129
? LENGTH: 2119
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8)..(8)
? OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-78129

```

Query Match	91.5%	Score 572;	DB 21;	Length 2119;
Best Local Similarity	90.6%	Pred. No. 2.6e-58;		
Matches 106; Conservative	6;	Mismatches 5;	Indels 0;	Gaps 0;

[illegible]

DQ
61 NNTMDWIQLEKIRAPSNVSTIHLILYPEAKGESVQPMKQDSL RVGEYVEACMAL 117

Db
307 NNSADWIQLERIRAPSNVSTIHLILYPEDAGGENVCQMKGQNL RVGEYVEACMAL 363

RESULT 5
US-09-724-676A-78129
; Sequence 78129, Application US/09724676A

```

1  APPLICANT: Compugen LTD
2  TITLE OF INVENTION: Variants of alternative splicing
3  FILE REFERENCE: 129181.4 Compugen
4  CURRENT APPLICATION NUMBER: US/09/724.676A
5  CURRENT FILING DATE: 2000-11-28
6  NUMBER OF SEQ ID NOS: 97222
7  SOFTWARE: patentln version 3.2
8  SEQ ID NO: 78129

```

OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-78129

Query Match	91.5%;	Score 572;	DB 21;	Length 2119;
Best Local Similarity	90.6%;	Pred. No. 2.6e-58;		
Matches 106; Conservative	6;	Mismatches 5;	Indels 0;	Gaps 0;

QY I EQCGIMHGNAVTFCEPYGPRELITTTCLNTTTSVLQPSIGSGCRFSISDPSITVSYAK 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 EQCGIMHGNAVTFCEPYGPRELITTTGLNTHASVLQPSIGSGCRFSISDPSITVLYAK 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY      61  NNTADMIQLEKIRAPSNVSTIYHILYIPEAKGSSVQOMKODSLRVGEYFEACNAL 117
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      307  NNSADMIQLEKIRAPSNVSTIHLIYIPEDAKGENVQOMKQENLRVGEYFEACNAL 363

```

```

RESULT 6
US-60-245-221-81
; Sequence 81, Application US/60245221
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00877
; CURRENT APPLICATION NUMBER: US/60/245, 221

```

```

:
: CURRENT FILING DATE: 2000-11-03
: NUMBER OF SEQ ID NOS: 234
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 81
: LENGTH: 2830
: TYPE: PRT
: ORGANISM: Human
US-60-245-221-81

```

Query Match	91.5%;	Score 572;	DB 37;	length 2830;
Best Local Similarity	90.6%;	Pred. No. 3.9e-58;		
Matches 106;	Conservative 6;	Mismatches 5;	Indels 0;	Gaps 0;

QY 1 EQCSTIMHGNVAVFCPEPYGPRELITTTCLNTTASVLQFSISGSCRRSYSDPSITVSFAK 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 142 EQCGAIMHGNAVFCEPYGPRELITTTGLNLTATSVLQFSISGSCRRSYSDPSIIVLYAK 201
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY      61  NNTADWIGLEKIRAPSNVSTIHLILPEEKAGESVQFQWKDOSTLRVEHYEACWAL 117
        ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      202  NNSADWIGLEKIRAPSNVSTIHLILPEADAGENVQFQWKDGNLRVGEHYEACWAL 258

```

RESULT 7
US-60-258-279-28

Sequence 28 Application US/60258279
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OR INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OR INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OR INVENTION: USES THEREOF
FILE REFERENCE: CLO01029-PROV
CURRENT APPLICATION NUMBER: US/60/258,279
CURRENT FILING DATE: 2000-12-27

```

: SOFTWARE: FastSQ for Windows Version 4.0
: SEQ ID NO 28
: LENGTH: 2830
: TYPE: PRS
: ORGANISM: HUMAN
US-60-258-279-28

```

Query Match	91.5%	Score 572;	DB 37;	Length 2830;
Best Local Similarity	90.6%	Pred. No. 3.9e-58;		
Matches 106; Conservative	6;	Mismatches 5;	Indels 0;	Gaps 0;

Oy

1 EOCGTIMHGNVAVTCPEPYGPRELITTCCLNTTASVTQFSISGSCRRSYSDSBITVSAAK 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db

142 EQCGAIMHGNAVTCPEPYGPRELITTGTLNTTASTAQFQSISGSCRRSYSDSIIVLYAK 201

```

Qy      61  NNTADWIOLEKIRAPSNVSTVIHILYLPEEAKGESVQFQWKDSL RVGEVYEACMAL 117
        ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      202  NNSADWIOLEKIRAPSNVSTIHLILYPEAKGSENVQFQWKQENL RVGEVYEACMAL 258

```

RESULT 8
US-09-949-003C-3671

```

1 Sequence 3671, Application US/09349003C
2
3 GENERAL INFORMATION:
4
5 APPLICANT: VENTER, J. Craig
6
7 TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCHIATRIC
8
9 TITLE OF INVENTION: DISORDERS, METHODS OF DETECTION AND USES THEREOF
10
11 FILE REFERENCE: C1000791
12
13 CURRENT APPLICATION NUMBER: US/09/349,003C
14
15 CURRENT FILING DATE: 2001-09-10

```

```

1 CURRENT APPLICATION NUMBER: US/09/949,003C
2 CURRENT FILING DATE: 2001-09-10
3 PRIOR APPLICATION NUMBER: 60/231,446
4 PRIOR FILING DATE: 2000-09-08
5 NUMBER OF SEQ ID NOS: 74065
6 SOFTWARE: PatentIn version 3.2
7 SEQ ID NO 3671
8 LENGTH: 2954
9 TYPE: PRT
10 ORGANISM: Homo sapiens
11 US-09-949-003C-3671

```

```
Query Match          91.5%; Score 572; DB 24; Length 2954;
Best Local Similarity 90.6%; Pred. No. 4.2e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVSXAK 60
DB 142 EOCGAIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVLVYAK 201

QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
DB 202 NNSADMIOLEKIRAPSNVSTIHIILYLPEDAKGENVQFQWKODSLRVEGYEACMAL 258

RESULT 9
US-10-170-205E-37739
; Sequence 37739, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37739
; LENGTH: 2992
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-37739

Query Match          91.5%; Score 572; DB 27; Length 2992;
Best Local Similarity 90.6%; Pred. No. 4.2e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVSXAK 60
DB 229 EOCGAIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVLVYAK 288

QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
DB 289 NNSADMIOLEKIRAPSNVSTIHIILYLPEDAKGENVQFQWKODSLRVEGYEACMAL 345

RESULT 10
US-60-453-050-10999
; Sequence 10999, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10999
; LENGTH: 2992
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-10999

Query Match          91.5%; Score 572; DB 37; Length 2992;
Best Local Similarity 90.6%; Pred. No. 4.2e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVSXAK 60
DB 229 EOCGAIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVLVYAK 288

QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
DB 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
```

```
DB 289 NNSADMIOLEKIRAPSNVSTIHIILYLPEDAKGENVQFQWKODSLRVEGYEACMAL 345

RESULT 11
US-60-453-135-10999
; Sequence 10999, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10999
; LENGTH: 2992
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-10999

Query Match          91.5%; Score 572; DB 37; Length 2992;
Best Local Similarity 90.6%; Pred. No. 4.2e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVSXAK 60
DB 229 EOCGAIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVLVYAK 288

QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
DB 289 NNSADMIOLEKIRAPSNVSTIHIILYLPEDAKGENVQFQWKODSLRVEGYEACMAL 345

RESULT 12
US-60-466-412-10999
; Sequence 10999, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10999
; LENGTH: 2992
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-10999

Query Match          91.5%; Score 572; DB 37; Length 2992;
Best Local Similarity 90.6%; Pred. No. 4.2e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVSXAK 60
DB 229 EOCGAIMHGNAAVTFCEPYGPRELITTCNTTTASVLOFSIGSGSCRFSSYSDPSITVLVYAK 288

QY 61 NNTADMIOLEKIRAPSNVSTVHIILYLPBEAKGESVQFQWKODSLRVEGYEACMAL 117
DB 289 NNSADMIOLEKIRAPSNVSTIHIILYLPEDAKGENVQFQWKODSLRVEGYEACMAL 345

RESULT 13
US-09-724-676-78128
; Sequence 78128, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
```

```
/ TITLE OF INVENTION: Variants of alternative splicing
/ FILE REFERENCE: 129181.4 Compugen
/ CURRENT APPLICATION NUMBER: US/09/724,676
/ CURRENT FILING DATE: 2000-11-28
/ NUMBER OF SEQ ID NOS: 97222
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 78128
/ LENGTH: 3107
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (8)..(8)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-78128

Query Match          91.5%; Score 572; DB 21; Length 3107;
Best Local Similarity 90.6%; Pred. No. 4.5e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAVTFCEPYGRELTTCTANTTASVLOFSIGSGSCRFPSYSDPSITVSYAK 60
DB 247 EOCGAIMHGNAVTFCEPYGRELTTGANTTTASVLOFSIGSGSCRFPSYSDPSITVLYAK 306

QY 61 NNTADWIOLEKIRAPSVNSTVTHILYLPBEAKGESVOFQWKODSLRGEVYEACMAL 117
DB 307 NNSADWIOLEKIRAPSVNSTIHIHLYLPBEAKGENVOFQWKQENLRVGEVYEACMAL 363

RESULT 14
US-09-724-676A-78128
/ Sequence 78128: Application US/09724676A
/ GENERAL INFORMATION:
/ APPLICANT: Compugen LTD
/ TITLE OF INVENTION: Variants of alternative splicing
/ FILE REFERENCE: 129181.4 Compugen
/ CURRENT APPLICATION NUMBER: US/09/724,676A
/ CURRENT FILING DATE: 2000-11-28
/ NUMBER OF SEQ ID NOS: 97222
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 78128
/ LENGTH: 3107
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (8)..(8)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-78128

Query Match          91.5%; Score 572; DB 21; Length 3107;
Best Local Similarity 90.6%; Pred. No. 4.5e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAVTFCEPYGRELTTCTANTTASVLOFSIGSGSCRFPSYSDPSITVSYAK 60
DB 247 EOCGAIMHGNAVTFCEPYGRELTTGANTTTASVLOFSIGSGSCRFPSYSDPSITVLYAK 306

QY 61 NNTADWIOLEKIRAPSVNSTVTHILYLPBEAKGESVOFQWKODSLRGEVYEACMAL 117
DB 307 NNSADWIOLEKIRAPSVNSTIHIHLYLPBEAKGENVOFQWKQENLRVGEVYEACMAL 363

RESULT 15
US-10-408-765-1624
/ Sequence 1624, Application US/10408765
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Bojin D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Bradford W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
```

```
/ APPLICANT: Warnock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ FILE REFERENCE: 660088.465
/ CURRENT APPLICATION NUMBER: US/10/408,765
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3025
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1624
/ LENGTH: 3427
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124,
/ LOCATION: 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
/ LOCATION: 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148,
/ LOCATION: 149, 150, 151, 152, 153, 154, 155, 156, 157, 158
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765-1624

Query Match          91.5%; Score 572; DB 30; Length 3427;
Best Local Similarity 90.6%; Pred. No. 5.1e-58;
Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EOCGTIMHGNAVTFCEPYGRELTTCTANTTASVLOFSIGSGSCRFPSYSDPSITVSYAK 60
DB 229 EOCGAIMHGNAVTFCEPYGRELTTGANTTTASVLOFSIGSGSCRFPSYSDPSITVLYAK 288

QY 61 NNTADWIOLEKIRAPSVNSTVTHILYLPBEAKGESVOFQWKODSLRGEVYEACMAL 117
DB 289 NNSADWIOLEKIRAPSVNSTIHIHLYLPBEAKGENVOFQWKQENLRVGEVYEACMAL 345
```

Search completed: August 4, 2005, 15:45:20
Job time : 495 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 15:26:18 ; Search time 39 Seconds
(without alignments)
288.650 Million cell updates/sec

Title: US-09-897-438b-2

Perfect score: 625

Sequence: 1 EDCGRIMGNNAVTFCEPYGP.....FQWKDSLRYGEVYEAQWAL 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 100 summaries

Database :

PIR 79: *
1: pirt: *
2: pirt2: *
3: pirt3: *
4: pirt4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	3461	2 S58870	reelin precursor -
2	81	13.0	640	1 ALASGR	glucan 1,4-alpha-g
3	81	13.0	640	2 A29166	glucan 1,4-alpha-g
4	77	12.3	301	2 T39401	probable ribose me
5	74	11.8	639	2 JT0479	glucan 1,4-alpha-g
6	73	11.7	639	2 J00607	glucan 1,4-alpha-g
7	72	11.5	316	2 G82204	transcription regu
8	71.5	11.4	297	2 T28775	hypothetical prote
9	70.5	11.3	868	2 A82515	conserved hypotet
10	70	11.2	1487	2 AG2560	hypothetical prote
11	69.5	11.1	528	2 S36520	li protein - human
12	69	11.0	460	2 C82399	conserved hypotet
13	68.5	11.0	611	2 S09500	dnak-type molecula
14	68.5	11.0	977	2 I52657	seizure-related pr
15	68	10.9	249	2 T31837	hypothetical prote
16	68	10.9	304	2 S04853	galactose 1-dehydr
17	68	10.9	462	2 G95233	Cof family protein
18	68	10.9	472	2 A99098	conserved hypotet
19	68	10.9	513	2 G96757	probable protein A
20	68	10.9	816	2 C69493	hypothetical prote
21	68	10.9	818	1 JC4397	peroxinectin precu
22	68	10.9	1074	2 G96504	probable En/Spm-11
23	67.5	10.8	3623	2 T08618	intrinsic factor-B
24	67	10.7	90	2 T01970	probable wound-ind
25	67	10.7	189	2 A82060	hypothetical prote
26	67	10.7	311	2 T23267	hypothetical prote
27	66.5	10.6	580	2 D69645	myo-inositol catab
28	66.5	10.6	195	2 T28773	hypothetical prote
29	66.5	10.6	1690	2 T31670	DNA-directed RNA p

30	66.5	10.6	1776	1 RRPWPM	genome polyprotein
31	66	10.6	366	2 T24546	hypothetical prote
32	66	10.6	373	2 B87887	protein T05F1.5 [1
33	66	10.6	2025	2 T21588	hypothetical prote
34	65.5	10.5	377	2 F82201	spermidine/putresc
35	65.5	10.5	503	2 S36549	li protein - human
36	65.5	10.5	540	1 A55145	chiamine-phosphate
37	65.5	10.5	549	2 T14606	probable sugar tra
38	65	10.4	112	2 T25554	hypothetical prote
39	65	10.4	359	2 H69287	hypothetical prote
40	65	10.4	417	2 H96733	hypothetical prote
41	65	10.4	422	2 A10145	probable substrate
42	65	10.4	451	2 F70743	hypothetical prote
43	65	10.4	565	1 HMIY87	hemagglutinin prec
44	65	10.4	1121	2 F86485	hypothetical prote
45	65	10.4	1187	1 A53661	protein-tyrosine-p
46	64.5	10.3	633	2 AE1344	hypothetical prote
47	64.5	10.3	747	2 T42599	minor capsid prote
48	64.5	10.3	867	1 GNLJSA	pol polyprotein -
49	64	10.2	440	2 T51278	hypothetical prote
50	64	10.2	674	2 AD3599	hemagglutinin limp
51	64	10.2	890	2 S44150	coat protein - str
52	64	10.2	1367	1 S48478	glucan 1,4-alpha-g
53	63.5	10.2	76	2 S19773	wound-induced prot
54	63.5	10.2	126	2 E83760	hypothetical prote
55	63.5	10.2	319	2 I51569	UVS.2 protein - Af
56	63.5	10.2	406	2 G87420	amino transferase,
57	63.5	10.2	467	2 AB1379	glutamate decarbox
58	63.5	10.2	573	2 T02743	lactase (EC 1.10.3
59	63.5	10.2	753	1 WZBBE8	gene 56 protein -
60	63.5	10.2	1967	2 S64604	hypothetical prote
61	63.5	10.2	2356	2 T27790	hypothetical prote
62	63	10.1	179	2 S30994	gene 49 protein -
63	63	10.1	376	2 D40511	hypothetical prote
64	63	10.1	441	2 C87215	probable lipoprote
65	63	10.1	456	2 C97296	UDP-N-acetylglucos
66	63	10.1	817	2 D85049	probable transposo
67	63	10.1	866	2 F88481	protein Cl6A3.1 [1
68	63	10.1	1797	2 T21889	hypothetical prote
69	63	10.1	1805	2 T21888	hypothetical prote
70	63	10.1	4377	2 A55575	ankyrin 3, long sp
71	63	10.1	4385	2 T29042	hypothetical prote
72	62.5	10.0	149	2 T25660	hypothetical prote
73	62.5	10.0	178	2 D72281	hypothetical prote
74	62.5	10.0	190	2 A98072	conserved hypotet
75	62.5	10.0	197	2 B89843	hypothetical prote
76	62.5	10.0	467	2 AC1748	glutamate decarbox
77	62.5	10.0	506	2 AH3252	dTMP kinase (EC 2.
78	62.5	10.0	567	2 T18462	hypothetical prote
79	62.5	10.0	585	2 S48950	hypothetical prote
80	62.5	10.0	687	2 T04927	probable serine/th
81	62.5	10.0	834	2 JQ1965	hypothetical 94k p
82	62.5	10.0	867	1 GNLJMP	pol polyprotein (c
83	62.5	10.0	913	2 D90183	ATP-dependent heli
84	62	9.9	211	2 A46458	human Cbl homolog
85	62	9.9	221	2 B70844	molycoprotein bios
86	62	9.9	365	2 T04247	hypothetical prote
87	62	9.9	380	2 T05638	hypothetical prote
88	62	9.9	421	2 T40614	G beta repeat prot
89	62	9.9	448	2 S03186	Ig heavy chain C r
90	62	9.9	506	2 A86169	hypothetical prote
91	62	9.9	535	2 S56261	probable membrane
92	62	9.9	955	2 B84645	probable villin 2
93	62	9.9	1215	2 T32734	myosin-1A - Acanth
94	62	9.9	1226	2 JC7503	protein-tyrosine-P
95	62	9.9	1254	2 G86379	protein F5A9.24 [1
96	62	9.9	1337	2 T41647	probable pre-mrna
97	62	9.9	2338	2 T25810	hypothetical prote
98	61.5	9.8	176	2 S76539	hypothetical prote
99	61.5	9.8	195	2 F71164	hypothetical prote
100	61.5	9.8	335	2 T28774	hypothetical prote

ALIGNMENTS

```

RESULT 1
S58870
reelin precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 21-Jul-2003
A:Accession: S58870, S71844, I49297
R:D'Arcangelo, G.; Mao, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.I.; Curran, T.
Nature 374, 719-723, 1995
A:Title: A protein related to extracellular matrix proteins deleted in the mouse mutant
A:Reference number: I49297; MUID:95231649; PMID:7715726
A:Accession: S58870
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-3461 <DAR>
A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487
R:D'Arcangelo, G.
submitted to the EMBL Data Library, April 1995
A:Reference number: S71844
A:Accession: S71844
A:Molecule type: mRNA
A:Residues: 1-115, 'T', 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>
A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-3461/Product: reelin #status predicted <MAT>
F:169-1795/Domain: EGF homology <EGF>

```

Query Match	100.0%	Score 625	DB 2	Length 3461
Best Local Similarity	100.0%	Pred. NO. 1.5e-55		
Matches 117	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	EQCGTMMGNVATCEPYPRELTTCNTTATTAATVLOPISGSGRFSYSDPSITVSYAK	60	
DB	230	EQCGTMMGNVATCEPYPRELTTCNTTATTAATVLOPISGSGRFSYSDPSITVSYAK	289	
QY	61	NNTADMTOLEKIRAPSNSTVIHILYLPEAKGSEVQOMKODSLRVGEVVEACAL	117	
DB	290	NNTADMTOLEKIRAPSNSTVIHILYLPEAKGSEVQOMKODSLRVGEVVEACAL	346	

RESULT 2
ALASCR
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Aspergillus niger*
N/Alternate names: 1,4-alpha-D-glucan glucohydrolase; glucoamylase
C/Species: *Aspergillus niger*
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: A90986; A91161; A05287; A22149; A25402
R/Boel, E.; Hansen, M.T.; Hjort, I.; Hoegh, I.; Fill, N.P.
EMBO J. 3, 1581-1585, 1984
A/Title: Two different types of intervening sequences in the glucoamylase gene from *Aspergillus niger*
Reference number: A90986; MUID:84261458; PMID:6204865

A:Cross-references: UNIPROT:P04064; GB:X00712; GB:K02466; NID:g2342; PIDDN:CMA25303.1; PI
A>Note: The authors translated the codon GAT for residue 317 as Asn
R:Svensson, B.; Larsen, K.; Gunnarsson, A.
Eur. J. Biochem. 154, 497-502, 1986
A>Title: Characterization of a glucosylase G2 from *Aspergillus niger*.
A:Reference number: A91161; MUD:66136085; PMID:3081341
A:Accession: comparison of forms G1 and G2
A:Accession: A91161
A:Molecule type: protein
A:Residues: 25-640 <SVE>
C:Comment: The large molecular form G1 is shown.
C:Comment: Smaller molecular forms of the enzyme, G2, arise by proteolytic cleavage(s) c
tive towards soluble poly- and oligosaccharides.

A:introns: 72/1, 167/3, 200/1, 412/3
C:superfamily: glucan 1,4-alpha-glucosylase with starch-binding domain; glucan 1,4-alpha-glucanase
Keywords: alternative splicing; extracellular protein; glycoprotein; glycosylase; hydrolase

Query Match	Score	DB 1	Length
F:1-24/Domain:signal sequence #status predicted <Sig>	13.04		640
E:23-447/Domain: glucan 1,4-alpha-glucosidase homology <GAG>			
F:25-660/Product: glucan 1,4-alpha-glucosidase G1 #status experimental <G1>			
F:23-556/Product: glucan 1,4-alpha-glucosidase G2, long form #status experimental <G3>			
F:23-556/Product: glucan 1,4-alpha-glucosidase G2, short form #status experimental <G32>			
F:195,419/Binding site: carbonyl:508,510,513,515,522,525,527,529,530,532/Binding site: carb			
F:465,467/Binding site: carbonyl:508,510,513,515,522,525,527,529,530,532/Binding site: carb			
F:476,486,488,496,499,500,501,502,504,506,512,514,517,518,520,524,526,528,531,534,535/Bit			

[illegible]

RESULT 3
A29166
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus awamori
N/Alternate names: 1,4-alpha-D-glucan glucosylhydrolase precursor; glucosylase precursor
C/Species: Aspergillus awamori
C/Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #ext_change 09-Jul-2004
C/Accession: A29166
R/Number: J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweikart, V.; Tai, Y.
Mol. Cell. Biol. 4, 2306-2315, 1984
A>Title: Molecular cloning and characterization of the glucosylase gene of Aspergillus
A/Reference number: A93066; PMID:85085934; PMID:6440004
A/Accession: A29166
A/Molecule type: DNA
A/Residues: 1-640 <NUN>
A/Cross-references: UNIPROT:P04064; GB:X02465; NID:G454405; PIDN:AAB59296.1; PID:G166505
C/Comment: See also PIR:ALASGR.
C/Genetic:
A/Introns: 72/1; 167/3; 200/1; 398/3
C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
C/Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hydroly
F.1-18/Domain: signal sequence #status predicted <SI>
F.19-640/Product: glucan 1,4-alpha-glucosidase #status predicted <MAT>
F.23-447/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

	Query Match	13.0%	Score 81;	DB 2;	Length 640;	
	Best Local Similarity	25.2%	Pred. No. 2.1,			
	Matches 34;	Conservative 15;	Mismatches 54;	Indels 32;	Gaps 4;	
Qy	9 GNAVTFEPPGPRELITTCINTTAAVLQFISGSGCRFSYS-----DPSITVSYAKN--	61				
Dd	497 GGITTTATPTFGSSGVSTSKTTAASKTSTSTSSCTTPPAVAVTFPLNTTYGENIY	556				
Qy	62 -----NTADWTGLEKIRAPSNVSIVIHILYLPEAKGSVOF-----Q	99				
Dd	557 LVGSIQLGDWMTSDGIALSDAKYTSSDDPLMYVTLP--AGSSEFKFIRESDDSVF	613				
Qy	100 WKQDSLAVGEVYLENC	114				
Dd	614 WESDPNREYTVPOAC	628				

RESULT 4
T39401
probable ribose methyltransferase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
!Accession: T39401


```

RESULT 8
T28775
hypothetical protein E03H12.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28775
R/Nelson, J.; Mohlmann, P.; Sansone, J.
submitted to the EMBL Data Library, June 1997
A/Description: The sequence of C. elegans cosmid E03H12.
A/Reference number: Z20520
A/Accession: T28775
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-297 <NBL>
A/Cross-references: UNIPROT:O02128; EMBL:AF000299; PIDD:MAC47978.1; GSPDB:GN00022; CESP
A/Experimental source: strain Bristol N2; clone E03H12
C/Genetics:
A:Gene: CESP:E03H12.2
A:Map position: 4
A:introns: 38/3; 91/3; 132/3; 170/1; 210/3; 261/2

```

Query Match	11.4%	Score 71.5	DB 2	Length 297
Best Local Similarity	21.7%	Pred. No. 8.2		
Matches 15, Conservative 16		Mismatches 21	Indels 17	Gaps 1

```

QY      10 NAYVFCEPIYPRELTT-----CLNTTASVLQFSLGSGSGRFSYSDP 52
       ::|||:::||:|||||
Db     45 SSCTLCDIYAVSKITQSNTYSNIQTAKIDSOLQCPKMTTNGTYTYTGSSNNYKLTFSDP 104

```

```
QY      53 SITVSYAKN 61
         |::|:
Db      105 SWTITYEKS 113
```

RESULT 9
A82515
conserved hypothetical protein XF2713 [imported] - *Xylella fastidiosa* (strain 9a5c)

C:/date: 18-Aug-2000 #sequence_ revision 20-Aug-2000 #text_change 09-Jul-2004
C:/Accession: A82515

R. Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequence Nature 406, 151-157, 2000

A1:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A1:Reference number: A82515; MUID:2035717; PMID:10910347
A1:Note: for a complete list of authors see reference number A59328 below

A;Accession: A82515
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-868 <SIM>

A/Cross-references: UNIPROT:O9P907; GB:AE004078; GB:AE003849; MID:95187832; PIDN:AAE57574
A/Experimental source: strain 945C
B/Simpson, A.J.G.; Reimach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrato, D.M.; Carter, H.; Cas-Neto, E.; Docena, C.; El-Dorry, H.; Falcinani, A.P.; Falcinella, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohnh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigry
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miraca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sasasaka
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328

A:Contents: annotation

C:Genetic8:

:Gene: XF2713

Query Match 11.3%; Score 70.5; DB 2; Length 866;
Best Local Similarity 29.1%; Pred. No. 38;
Matches 16; Conservative 11; Mismatches 21; Gaps 1.

19 GPRELTTCLNTTAS-----VLQFSIGSGCRFSYSDPSITVSYAKNNTADW 66

Db 372 GPKVLTQTGTTKATGSKDPEYIMKFLQEGGYRPNYDGTHTGVITYDNNNASNW 426

RESULT 10
AG2560

A1:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C1:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

R. Kaneo, T.; Nakamura, Y.; Wolk, C. P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguchi, S.; Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. *Nat. Res.* 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2560
A;Status: Preliminary

A;Residues: 1-1487 <KUR>
A;Cross-references: UNIPROT:O8YV40, GB:AP003603, PIDD:BAE77408.1, PIDD:G17134851, GSPDB:GN
A;Experimental source: Strain PCC 7120

A; Genome: plasmid

Query Match	11.2%	Score 70;	DB 2;	Length 1487;
Best Local Similarity	29.6%	Pred. No. 83;		
Matches 24;	Conservative 12;	Mismatches 27;	Indels 18;	Gaps 4;

```
Oy      5 TIMH-----GNAVTFCEPYGPBRLTTCNTTTASVLQFSGSSCRFSYD-PSIT 55  
||:| ||: | :| :| :| :|  
Db      471 TITHSNTINNVGAIALNSQ-VSTTEITNNLNNIT-----GTGLLLTNVSDTSISIT 521
```

```
QY      56 VSYAKNTADWIOLEKIRAPS 76
          : | | | | : : : :
Db      522 NNTISNTTGDGIQLSQVISP 542
```

RESULT 11
S36520

C:Species: human papillomavirus type 34
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S36520

Ridellius, H.; Hoermann, B.
submitted to the EMBL Data Library, August 1993
A>Description: Primer-directed sequencing of human papillomavirus types
A;Reference number: S36469

A/Accession: S36520
A/Molecule type: DNA
A/Residues: 1-528
A/Cross-references: UNIPROT:P36738, EMBL:X74476, NID:G396989, PIDN:CAA52560.1, PID:G39695
A/Superfamily: papillomavirus L1 protein

Query Match	11.1%	Score 69.5	DB 2	Length 528
Best Local Similarity	22.1%	Pred. No. 27		
Matches 27; Conservative	20	Mismatches 36	Indels 39	Gaps 4

Oy 4 GTTTHGNAVTFCEPY-----GPRLTTTCNLTNTTASVLQFSIGSGSCRFSYS 50
 | | | | | | | | |
Db 320 GSNWSSDAQIENRPEYWLAKAOGNGICMHNQLFLTVADITRTSRNSVCVGTO----- 373

```
QY      51 DPGITVSYAKNNTADMI-----OLEKTRAPSNSTVIHIIYLPEAKGESVOF 98  
          ||| | : ::||| |::| |:  
Db     374 -TSTLPDYANSFKKEYLRHAERYDLOFPOLCKINLTVDVTIYH-----SSMSGLE 422
```

QY	99	QW	100
Db	426	QW	427

RESULT 12
C82399
conserved hypothetical protein VCA0931 [imported] - Vibrio cholerae (strain N16961 serog
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: C82399
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, F
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
A/Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:2040683; PMID:10952301
A/Accession: C82399
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-460 <HEI>
A/Cross-references: UNIPROT:Q9KLI6; GB:AE004420; GB:AE003853; NID:96658361; PIDN:AAF9682
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VCA0931
A/Map position: 2

Query Match 11.0%; Score 69; DB 2; Length 460;
Best Local Similarity 23.4%; Pred. No. 25;
Matches 26; Conservative 18; Mismatches 43; Indels 24; Gaps 3;

Qy 10 NAVTFCPPYGPRLT---TTCANTTASYLQPSIGSGCRFSYSDPSITVSYAKNTAD 65
Db 70 NAARICELYATDILTFKRRKYVGTSLSSVINFTVKNYTGSEOSWTEKILTTIDILKNGND 129
66 WID-----LEKIRAPSNVSTVTHL-----YLPEAKGESV 96
Qy 130 YAOELIOTRCTRGADVARKEURFSEAVAGQIHSLDEHWNNGGRPEQRKEAI 180

RESULT 13
S09500
dnaK-type molecular chaperone dnaK - Bacillus subtilis
N/Alternate names: heat shock protein
C/Species: Bacillus subtilis
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C/Accession: S09500; S27505; S06596; H69617
R/Wetzstein, M.; Dedio, J.; Schumann, W.
A/Title: Complete nucleotide sequence of the Bacillus subtilis dnaK gene.
A/Reference number: S09500; MUID:90245667; PMID:2110662
A/Accession: S09500
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-611 <MET>
A/Cross-references: UNIPROT:P17820; EMBL:X52064; NID:939888; PIDN:CAA36286.1; PID:958085
R/Wetzstein, M.; Voelker, U.; Dedio, J.; Loebau, S.; Zuber, U.; Schlesswohl, M.; Hergert,
submitted to the EMBL Data Library, January 1992
A/Description: Cloning, sequencing, and molecular analysis of the dnaK locus from Bacill
A/Reference number: S27503
A/Accession: S27505
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-611 <ME2>
A/Cross-references: EMBL:M84964; NID:9143056; PIDN:AAA22528.1; PID:9143059
R/Hearne, C.M.; Elar, D.J.
Nucleic Acids Res. 17, 8373, 1989
A/Title: Nucleotide sequence of a Bacillus subtilis gene homologous to the dnaK gene of
A/Reference number: S06596; MUID:90045965; PMID:2510131
A/Accession: S06596
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-98, 'R', 100-149, 'P', 165, 'L', 167-214 <HEA>
A/Cross-references: EMBL:X16393
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brouillet, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Koningsberg, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapides, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mueel,
Y.M.; Ogawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Rodhe, B.; Rose, M.; Sadale, Y.; Sato, T.; Scallan,
A/Authors: Schlecht, S.; Schreier, R.; Scottone, P.; Sekiguchi, J.; Sekowska, A.; Seror,
akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: H69617
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-611 <KUN>
A/Cross-references: GB:Z99117; GB:AL009126; NID:92634966; PIDN:CA814489.1; PID:92634993
A/Experimental source: strain 168
C/Genetics:
A/Gene: dnaK
A/Start codon: GNG
C/Function:
A/Description: Involved in protein folding and assembling/disassembling of protein complex
C/Superfamily: heat shock protein 70
C/Keywords: ATP; molecular chaperone

Query Match 11.0%; Score 68.5; DB 2; Length 611;
Best Local Similarity 25.2%; Pred. No. 40;
Matches 26; Conservative 20; Mismatches 40; Indels 17; Gaps 5;

Qy 3 CGTIMGNNAVTFCEPYGPRLTTCANTTASYLQF-----SIGSGCRFSYSDPSITVS 57
Db 15 CVALEGE-----PKYIANAEGNRITPSVAFRNGRGROYGEVAKRSINPNITWS 66

Qy 58 YAKNTADW-IQLE-KIRAPSNVSTV-I-HILYLPPEAKGESV 96
Db 67 IKRMGTGDKVKEIKDYTPQEVSAIILQHLKSYAESYLGRTV 109

RESULT 14
I52657
seizure-related protein SEZ-6 precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C/Accession: I52657
R/Shimizu-Nishikawa, K.; Kajiwara, K.; Kimura, M.; Katsuki, M.; Sugaya, E.
Brain Res. Mol. Brain Res. 28, 201-210, 1995
A/Title: Cloning and expression of SEZ-6, a brain-specific and seizure-related cDNA.
A/Reference number: I52657; MUID:95240392; PMID:7723619
A/Accession: I52657
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-977 <RES>
A/Cross-references: UNIPROT:Q62269; GB:D29763; NID:9693909; PIDN:BAA06167.1; PID:9693910
F/354-409/Domain: complement factor H repeat homology <FH01>
F/529-586/Domain: complement factor H repeat homology <FH02>
F/707-762/Domain: complement factor H repeat homology <FH03>
F/768-827/Domain: complement factor H repeat homology <FH04>
F/835-892/Domain: complement factor H repeat homology <FH05>

Query Match 11.0%; Score 68.5; DB 2; Length 977;
Best Local Similarity 28.3%; Pred. No. 71;
Matches 26; Conservative 14; Mismatches 35; Indels 17; Gaps 3;

Qy 19 GPRELTTTCANTTASYLQPSIGSGCRFSYSDPSITVSYAKNTADWIOLEKIRAPSNV 78
Db 218 GDBETTTTITTTTIVTVQ---PPGCSNWFSGREGS-----LDSPTABSSP 261

Qy 79 STV-IHILYLPPEAKGESVQPKKQSLRNGE 109
Db 262 SDVGLDFYIISVYPPGVGAIKENISLQGE 293

RESULT 15

T31837
 hypothetical protein T05B4.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T31837
 R:Bradehaw, H.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid T05B4.
 A:Reference number: Z21092
 A:Accession: T31837
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-249 <BRA>
 A:Cross-references: UNIPROT:O16420; EMBL:AF016445; PIDN:AAC69060.1; GSPDB:GN00023; CESP:
 A:Experimental source: strain Bristol N2; clone T05B4
 C:Genetics:
 A:Gene: CESP.T05B4.11
 A:Map position: 5
 A:introns: 34/3; 121/1
 C:Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 10.9%; Score 68; DB 2; Length 249;
 Best Local Similarity 28.6%; Pred. No. 15;
 Matches 22; Conservative 15; Mismatches 26; Indels 14; Gaps 4;

QY 9 GNAVTCPEYGPRLFTT-----CLNTTASVLQFSIGSGCRPSYSDPSIT-VSYA 59
 DB 167 GNDLSICQNIQMGQSFVNTYCQKSCGRCPSITTSQ---YNNGGSGCTSYVADSSSNCASWA 223
 QY 60 KNN--TADWIQLEKIRA 74
 DB 224 KNGFCTNPFYVDQRRRA 240

Search completed: August 4, 2005, 15:36:13
 Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 4, 2005, 15:25:38 / Search time 167 Seconds
(without alignments)
358.762 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 625
Sequence: 1 EQCGTIMGNAVTFCBPYGP.....FMWKDSLRYGEVYEAACMAL 117

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database: UniProt_03:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	1181	2	Q8C4E4
2	625	100.0	3461	1	RELN_MOUSE
3	596	95.4	3462	1	RELN_RAT
4	572	91.5	3460	1	RELN_HUMAN
5	570	91.2	216	2	Q6Q145
6	418	66.9	3209	1	RELN_CHICK
7	223	35.7	279	2	Q8C978
8	118	18.9	52	1	RELN_BOVIN
9	110.5	17.7	3008	2	Q6Q144
10	106.5	17.0	747	2	Q64FW1
11	89.5	14.3	338	2	Q93576
12	88.5	14.2	205	2	Q6PX75
13	87.5	14.0	308	2	Q8AYR0
14	87.5	14.0	635	2	Q93575
15	81	13.0	639	2	Q870C8
16	81	13.0	640	1	AMYG_ASPNG
17	80.5	12.9	291	2	Q8AY69
18	80	12.8	640	2	Q6DNH5
19	80	12.8	726	2	Q6YFE6
20	79	12.6	636	2	Q6CUT9
21	79	12.6	1078	2	Q8AKX7
22	77	12.3	301	2	Q94631
23	77	12.3	640	2	Q6DUV5
24	75	12.0	1012	2	Q6BNF2
25	75	12.0	433	2	Q8AXK6
26	74	11.8	264	2	Q6ACQ4
27	74	11.8	639	1	AMYG_ASPKA
28	73.5	11.8	1130	2	Q88262
29	73	11.7	639	1	AMYG_ASPSH
30	72	11.5	316	2	Q9K569
31	72	11.5	639	2	Q12537

32	71.5	11.4	297	2	Q02128	Q02128 caenorhabdi
33	71.5	11.4	1178	2	Q7UPT4	Q7UPT4 rhodopirell
34	71	11.4	426	2	Q7MHP3	Q7MHP3 vibrio vuln
35	71	11.4	426	2	Q8DC71	Q8DC71 vibrio vuln
36	71	11.4	639	2	Q76L97	Q76L97 aspergillus
37	71	11.4	869	2	Q8EDY5	Q8EDY5 shewanella
38	71	11.4	1201	2	Q76G29	Q76G29 zea mays (m
39	70.5	11.3	271	2	Q8PIY4	Q8PIY4 streptococc
40	70.5	11.3	271	2	Q9A0X2	Q9A0X2 streptococc
41	70.5	11.3	271	2	Q8K887	Q8K887 streptococc
42	70.5	11.3	347	2	Q735K0	Q735K0 bacillus ce
43	70.5	11.3	403	2	Q8RK83	Q8RK83 mycobacteri
44	70.5	11.3	868	2	Q879X4	Q879X4 xylella fas
45	70.5	11.3	868	2	Q9PA07	Q9PA07 xylella fas
46	70	11.2	123	2	Q18797	Q18797 gorilla gor
47	70	11.2	809	1	ILAR_HORSE	Q6WQ24 equus cabal
48	70	11.2	1487	2	Q8YK40	Q8YK40 anabaena sp
49	69.5	11.1	125	2	Q66072	Q66072 canine hecp
50	69.5	11.1	313	2	Q88VP3	Q88VP3 lactobacill
51	69.5	11.1	528	1	VL1_HP34	Q36738 human papil
52	69.5	11.1	546	2	Q8FQX5	Q8FQX5 corynebacte
53	69.5	11.1	605	2	Q7TFC9	Q7TFC9 mus musculu
54	69.5	11.1	667	2	Q9M3P4	Q9M3P4 drosophila
55	69.5	11.1	991	2	Q7T8K2	Q7T8K2 mus musculu
56	69	11.0	364	1	DFRA_CALCH	Q759X3 asbhyra goss
57	69	11.0	417	2	Q759X3	Q9K118 vibrio chol
58	69	11.0	460	2	Q9K118	Q6C777 yarrowia li
59	69	11.0	1257	2	Q6C7N7	Q6CWF5 kluyveromyc
60	68.5	11.0	210	2	Q6CWF5	Q7F592 oryza sativ
61	68.5	11.0	533	2	Q7F592	Q62224 mus musculu
62	68.5	11.0	605	2	Q62224	Q17820 bacillus su
63	68.5	11.0	610	1	DNAB_BACSU	Q62269 mus musculu
64	68.5	11.0	977	2	Q62269	Q62223 mus musculu
65	68.5	11.0	991	2	Q62223	Q16420 caenorhabdi
66	68	10.9	249	2	Q16420	Q11886 pseudomonas
67	68	10.9	304	1	GAL_PSEFL	Q8C014 mus musculu
68	68	10.9	383	2	Q93885	Q93885 emeritella
69	68	10.9	386	2	Q93885	Q94783 mus musculu
70	68	10.9	404	2	Q9D7F3	Q9D7F3 mus musculu
71	68	10.9	427	1	YVF5_BRAJA	Q7XR19 oryza sativ
72	68	10.9	448	2	Q7XR19	Q97NM3 streptococc
73	68	10.9	462	2	Q97NM3	Q8DNC5 streptococc
74	68	10.9	472	2	Q8DNC5	Q8DNC5 mus musculu
75	68	10.9	497	2	Q8BYL2	Q8BYL2 mus musculu
76	68	10.9	499	2	Q922P4	Q81973 arabidopsis
77	68	10.9	512	2	Q8L673	Q9CAT3 arabidopsis
78	68	10.9	513	2	Q9CAT3	Q8BMD5 mus musculu
79	68	10.9	523	2	Q8BMD5	Q8BYN1 mus musculu
80	68	10.9	816	2	Q28331	Q28331 archaeoglob
81	68	10.9	816	2	Q28331	Q26059 pacifastacu
82	68	10.9	818	2	Q26059	Q69X06 oryza sativ
83	68	10.9	826	2	Q69X06	Q6ZB12 oryza sativ
84	68	10.9	884	2	Q6ZB12	Q91NY8 arabidopsis
85	68	10.9	1001	2	Q91NY8	Q91NY8 arabidopsis
86	68	10.9	1074	2	Q9C6X8	Q9C6X8 arabidopsis
87	68	10.9	2281	2	Q64HX2	Q64HX2 sapovirus c
88	67.5	10.8	427	2	Q74682	Q74682 aspergillus
89	67.5	10.8	3623	2	Q70244	Q70244 ratius norv
90	67	10.7	90	2	Q82615	Q82615 arabidopsis
91	67	10.7	147	2	Q6CM12	Q6CM12 yarrowia li
92	67	10.7	179	2	Q6CM12	Q6M142 bdellovibri
93	67	10.7	189	2	Q9KP16	Q9KP16 vibrio chol
94	67	10.7	236	2	Q73BM8	Q73BM8 bacillus ce
95	67	10.7	236	2	Q9BZ35	Q9BZ35 homo sapien
96	67	10.7	251	2	Q6NSB0	Q6NSB0 homo sapien
97	67	10.7	331	2	Q21182	Q21182 caenorhabdi
98	67	10.7	497	2	Q6PHA8	Q6PHA8 mus musculu
99	67	10.7	514	2	Q6BWA3	Q6BWA3 homo sapien
100	67	10.7	580	1	IOLD_BACSU	P42415 bacillus su

ALIGNMENTS

```

RESULT 1
Q8C4E4 PRELIMINARY; PRT; 1181 AA.
ID Q8C4E4
AC Q8C4E4;
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched
DE library; clone: C230051M12 product: reelin, full insert sequence.
GN Name=Rein;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA The FANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishii K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Yoneda Y., Ishikawa T., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,

```

```

RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBCP databases.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
DR EMBL; AK082447; BAC38493.1; -.
DR MGD; MG1:103022; ReIn.
DR GO; GO:0055783; C:extracellular matrix (sensu Metazoa); TAS.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IMP.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0016477; P:cell migration; IMP.
DR GO; GO:0009094; P:cellular morphogenesis during differentiation; IMP.
DR GO; GO:0007529; P:establishment of synaptic specificity at ne...; IMP.
DR GO; GO:0010001; P:glial cell differentiation; IMP.
DR GO; GO:0051057; P:positive regulation of small GTPase mediate...; IDA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002860; Glyco_hydro_BNR.
DR InterPro; IPR002661; Reeler.
DR InterPro; IPR011040; Stailidase.
DR Pfam; PF02012; BNR; 4.
DR Pfam; PF02014; Reeler; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
KW EGF-like domain.
SQ SEQUENCE 1181 AA; 131491 MW; 3CEA91P9EF72C720 CRC64;
Query Match 100.0%; Score 625; DB 2; Length 1181;
Best Local Similarity 100.0%; Pred. No. 3,76-58;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ECGTIMGNNAVTCBPYGPRLTTTCNTTASVLAQSGSGCRFSYSDPSITVSAYK 60
DB 230 ECGTIMGNNAVTCBPYGPRLTTTCNTTASVLAQSGSGCRFSYSDPSITVSAYK 289
QY 61 NNTADMIOLEKTRASNSVTHIILYLPBEAKGSEVOGQKODSLRGEVYEAACAL 117
DB 290 NNTADMIOLEKTRASNSVTHIILYLPBEAKGSEVOGQKODSLRGEVYEAACAL 346
RESULT 2
REIN_MOUSE STANDARD; PRT; 3461 AA.
ID REIN_MOUSE
AC Q60841; Q9CUA6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Reelin precursor (EC 3.4.21.-) (Reeler protein).
GN Name=Rein; Synonyms=RL;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=95231649; PubMed=7715726; DOI=10.1038/374719a0;
RA D'Arcangelo G., Mao G.G., Chen S.-C., Soares H.D., Morgan J.I.,
RA Curran T.;
RT "A protein related to extracellular matrix proteins deleted in the
RT mouse mutant reeler.";
RL Nature 374:719-723(1995).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98086481; PubMed=9411911; DOI=10.1006/geno.1997.4983;
RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demitov D.,
RA Goffinet A.M.;
RT "Genomic organization of the mouse reelin gene.";
RL Genomics 46:240-250(1997).
RN [3]
RP SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).
RC STRAIN=BA6B/c; TISSUE=Brain;
RX MEDLINE=95375789; PubMed=7647795;

```

RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,
 RA Ohashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,
 RA Nako K., Katsuki M., Hayashizaki Y.;
 RT "The reeler gene encodes a protein with an EGF-like motif expressed by
 RT pioneer neurons.";
 RT Nat. Genet. 10:77-83(1995).
 RN [4]
 RP SEQUENCE OF 3044-3461 FROM N.A. (ISOFORM 2).
 RC MEDLINE=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikido I., Oseto N., Saito R., Suzuki H., Yamana H., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.W.,
 RA Schmitt L.M., Kanpin A., Matsuda H., Baralov S., Beisel K.W.,
 RA Blake J.M., Brad D., Brasic V., Chotina C., Corbani L.B., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada M.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmig L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirotsune-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arkawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RT Nature 420:563-573(2002).
 RN [5]
 RP CHARACTERIZATION.
 RC MEDLINE=97141547; PubMed=8987733;
 RX MEDLINE=97141547; PubMed=8987733;
 RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,
 RA Curran T.;
 RT "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal
 RT antibody.";
 RT J. Neurosci. 17:23-31(1997).
 RN [6]
 RP CHARACTERIZATION.
 RC MEDLINE=21634904; PubMed=11689558; DOI=10.1074/jbc.M106996200;
 RX MEDLINE=21634904; PubMed=11689558; DOI=10.1074/jbc.M106996200;
 RA Quattrocchi C.C., Wamnes F., Pericak A.M., Claret S.A.,
 RA D'Arcangelo G., Farace M.G., Keller F.;
 RT "Reelin is a serine protease of the extracellular matrix.";
 RT J. Biol. Chem. 277:303-309(2002).
 RN [7]
 RP TISSUE SPECIFICITY.
 RC MEDLINE=97325946; PubMed=9182958;
 RX MEDLINE=97325946; PubMed=9182958;
 RA Schiffmann S.N., Bernier B., Goffinet A.M.;
 RT "Reelin mRNA expression during mouse brain development.";
 RT Eur. J. Neurosci. 9:1055-1071(1997).
 RN [8]
 RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC MEDLINE=99263456; PubMed=10328932; DOI=10.1006/exnr.1999.7019;
 RX MEDLINE=99263456; PubMed=10328932; DOI=10.1006/exnr.1999.7019;
 RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergeyck V.,
 RA Goffinet A.M.;
 RT "Evolutionarily conserved, alternative splicing of reelin during brain
 RT development.";
 RT Exp. Neurol. 156:229-238(1999).
 RN [9]
 RP BINDING TO VLDLR AND APOER2.
 RC MEDLINE=20036019; PubMed=10571241; DOI=10.1016/S0896-6273(00)80861-2;
 RX MEDLINE=20036019; PubMed=10571241; DOI=10.1016/S0896-6273(00)80861-2;
 RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumbly M.C.,
 RA Cooper J.A., Herz J.;

RT "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
 RT tyrosine phosphorylation of disabled-1 and modulates tau
 RT phosphorylation.";
 RT Neuron 24:481-489(1999).
 RN [10]
 RP FUNCTION.
 RC MEDLINE=20359755; PubMed=10880573; DOI=10.1073/pnas.150040497;
 RX Yip J.W., Yip Y.P.L., Nakajima K., Capriotti C.;
 RT "Reelin controls position of autonomic neurons in the spinal cord.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).
 RN [11]
 RP FUNCTION.
 CC -1- FUNCTION: Extracellular matrix serine protease that plays a role
 CC in layering of neurons in the cerebral cortex and cerebellum.
 CC Regulates microtubule function in neurons and neuronal migration.
 CC Affects migration of sympathetic preganglionic neurons in the
 CC spinal cord, where it seems to act as a barrier to neuronal
 CC migration. Enzymatic activity is important for the modulation of
 CC cell adhesion. Binding to the extracellular domains of lipoprotein
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
 CC Dab1 and modulation of Tau phosphorylation.
 CC -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: Secreted.
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q60841-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q60841-2; Sequence=VSP_005577;
 CC Name=3;
 CC IsoId=Q60841-3; Sequence=VSP_005578;
 CC -1- TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is
 CC abundantly produced during brain ontogenesis by the Cajal-Retzius
 CC cells and other pioneer neurons located in the telencephalic
 CC marginal zone and by granule cells of the external granular layer
 CC of the cerebellum. Expression is located in deeper layers in the
 CC developing hippocampus and olfactory bulb, low levels of
 CC expression are also detected in the immature striatum. At early
 CC developmental stages, expressed also in hypothalamic
 CC differentiation fields, tectum and spinal cord. A moderate to low
 CC level of expression occurs in the septal area, striatal fields,
 CC habenular nuclei, some thalamic nuclei, particularly the lateral
 CC geniculate, the retina and some nuclei of the reticular formation
 CC in the central field of the medulla. Very low levels found in
 CC liver and kidney. No expression in radial glial cells, cortical
 CC plate, Purkinje cells and inferior olivary neurons. The minor
 CC isoform 2 is only expressed in non neuronal cells. The minor
 CC isoform 3 is found in the same cells as isoform 1, but is almost
 CC undetectable in retina and brain stem.
 CC -1- DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.
 CC Expression increases up to birth and remains high from post-natal
 CC day 2 to 11 in both cerebellum and fore/midbrain. Expression
 CC declines thereafter and is largely brain specific in the adult.
 CC DOMAIN: The basic C-terminal region is essential for secretion.
 CC -1- PTM: N-glycosylated and to a lesser extent also O-glycosylated.
 CC -1- DISEASE: Defects in Reelin are the cause of the autosomal recessive
 CC reeler (rl) phenotype which is characterized by impaired motor
 CC coordination, tremors and ataxia. Neurons in affected mice fail to
 CC reach their correct locations in the developing brain, disrupting
 CC the organization of the cerebellar and cerebral cortices and other
 CC laminated regions.
 CC -1- SIMILARITY: Belongs to the reelin family.
 CC -1- SIMILARITY: Contains 15 BNR repeats.
 CC -1- SIMILARITY: Contains 8 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 reelin domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U24703; AAB91599.1; -

DR EMBL, D63520; BAA09788.1; ALT_INIT.
 DR EMBL, AK017094; BAB30592.1; -.
 DR MGD; MG1:103022; Reln.
 DR GO; GO:0005615; C:extracellular space; IDA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IMP.
 DR GO; GO:0007420; P:brain development; IMP.
 DR GO; GO:0016477; P:cell migration; IMP.
 DR GO; GO:0000904; P:cellular morphogenesis during differentiation; IMP.
 DR GO; GO:0007529; P:establishment of synaptic specificity at ne. . .; IMP.
 DR GO; GO:0010001; P:glial cell differentiation; IMP.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR002860; Glyco_hydro_BNR.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR002861; Reeler.
 DR InterPro; IPR011040; Stailidase.
 DR Pfam; PF02012; BNR; 15.
 DR Pfam; PF00008; EGF; 3.
 DR Pfam; PF02014; Reeler; 1.
 DR SMART; SM00181; EGF; 5.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 6.
 DR PROSITE; PS50026; EGF_3; 5.

Query Match 100.0%; Score 625; DB 1; Length 3461;
 Best Local Similarity 100.0%; Pred.No. 1.3e-57;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ECGGTMGNAATPCGYPRELTTCGLNTTASVLOFSGSGCRFSYSDPSITVSYAK 60
 Db 230 ECGGTMGNAATPCGYPRELTTCGLNTTASVLOFSGSGCRFSYSDPSITVSYAK 289
 Oy 61 NNTADMIQLEKIRAPSNVSTVTHILVLPKAKGESVQFQWKDLSLVGEVYKCMWL 117
 Db 290 NNTADMIQLEKIRAPSNVSTVTHILVLPKAKGESVQFQWKDLSLVGEVYKCMWL 346

RESULT 3
 REIN RAT STANDARD; PRT; 3462 AA.
 AC P58751; O80765;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Reelin precursor (EC 3.4.21.-).
 GN Name=Rein;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Cerebellum;
 RA Kikkawa S., Terashima T.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A., AND DISEASE.
 RC TISSUE=Cerebellum;
 RX MEDLINE=22557166; PubMed=12670697; DOI=10.1016/S0169-328X(02)00650-2;
 RA Yokoi N., Namee M., Wang H.-W., Kojima K., Fuse M., Yasuda K.,
 RA Serikawa T., Seino S., Komeda K.;
 RT "Rat neurological disease creeping is caused by a mutation in the
 RT reelin gene";
 RL Reelin Res. Mol. Brain Res. 112:1-7(2003).
 RN [3]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=9263436; PubMed=10328932; DOI=10.1006/exnr.1999.7019;
 RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergsyck V.,
 RA Goffinet A.M.;
 RT "Evolutionarily conserved, alternative splicing of reelin during brain
 RT development";
 RL Exp. Neurol. 156:229-238(1999).
 CC -!- FUNCTION: Extracellular matrix serine protease that plays a role

CC in layering of neurons in the cerebral cortex and cerebellum.
 CC Regulates microtubule function in neurons and neuronal migration.
 CC Affects migration of sympathetic preganglionic neurons in the
 CC spinal cord, where it seems to act as a barrier to neuronal
 CC migration. Enzymatic activity is important for the modulation of
 CC cell adhesion. Binding to the extracellular domains of lipoprotein
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
 CC Dab1 and modulation of Tau phosphorylation (By similarity).
 CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P58751-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P58751-2; Sequence=VSP_005579;
 CC Name=3;
 CC IsoId=P58751-3; Sequence=VSP_005580;
 CC -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
 CC by the Cajal-Retzius cells and other pioneer neurons located in
 CC the telencephalic marginal zone and by granule cells of the
 CC external granular layer of the cerebellum.
 CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By
 CC similarity).
 CC -!- DISEASE: Defects in Reelin are the cause of the creeping phenotype,
 CC which is characterized by tremor, gait ataxia, cerebellar
 CC hypoplasia and abnormal neuronal migration (particularly in the
 CC cerebral cortex and hippocampus). The mutation is due to a
 CC nucleotide insertion at codon 1892 which results in a
 CC translational frameshift and truncation of the protein.
 CC -!- SIMILARITY: Belongs to the reelin family.
 CC -!- SIMILARITY: Contains 15 BNR repeats.
 CC -!- SIMILARITY: Contains 8 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 reelin domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL, AB049473; BAB78470.1; -.
 CC DR EMBL, AB062680; BAC75467.1; -.
 CC DR HSSP; P05106; ILSG.
 CC DR RGD; 3553; Reln.
 CC DR GO; GO:0001764; P:neuronal migration; IMP.
 CC DR InterPro; IPR000742; EGF_2.
 CC DR InterPro; IPR006209; EGF_like.
 CC DR InterPro; IPR002860; Glyco_hydro_BNR.
 CC DR InterPro; IPR006210; IEGF.
 CC DR InterPro; IPR002861; Reeler.
 CC DR InterPro; IPR011040; Stailidase.
 CC DR Pfam; PF02012; BNR; 15.
 CC DR Pfam; PF00008; EGF; 3.
 CC DR Pfam; PF02014; Reeler; 1.
 CC DR SMART; SM00181; EGF; 6.
 CC DR PROSITE; PS00022; EGF_1; 7.
 CC DR PROSITE; PS01186; EGF_2; 6.
 CC DR PROSITE; PS50026; EGF_3; 5.
 CC DR PROSITE; PS51019; REELIN; 1.
 CC KW Alternative splicing; Cell adhesion; Developmental protein;
 CC EGF-like domain; Extracellular matrix; Glycoprotein; Hydrolase;
 CC Repeat; Serine protease; Signal.
 CC KM
 CC FT STGNAL 1 27 Potential.
 CC FT CHAIN 28 3462 Reelin.
 CC FT DOMAIN 28 192 Reelin.
 CC FT DOMAIN 672 703 EGF-like 1.
 CC FT DOMAIN 1031 1062 EGF-like 2.
 CC FT DOMAIN 1410 1443 EGF-like 3.
 CC FT DOMAIN 1766 1797 EGF-like 4.

```

FT DOMAIN 2130 2162 EGF-like 5.
FT DOMAIN 2479 2510 EGF-like 6.
FT DOMAIN 2854 2885 EGF-like 7.
FT DOMAIN 3229 3261 EGF-like 8.
FT REPEAT 594 605 BNR 1.
FT REPEAT 800 811 BNR 2.
FT REPEAT 953 964 BNR 3.
FT REPEAT 1158 1169 BNR 4.
FT REPEAT 1324 1335 BNR 5.
FT REPEAT 1536 1547 BNR 6.
FT REPEAT 1687 1698 BNR 7.
FT REPEAT 1885 1896 BNR 8.
FT REPEAT 2044 2055 BNR 9.
FT REPEAT 2251 2262 BNR 10.
FT REPEAT 2400 2411 BNR 11.
FT REPEAT 2599 2610 BNR 12.
FT REPEAT 2779 2790 BNR 13.
FT REPEAT 2980 2991 BNR 14.
FT REPEAT 3364 3375 BNR 15.
FT DOMAIN 3433 3462 Arg-rich (basic).
FT CARBOHYD 142 142 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 259 259 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 291 291 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 307 307 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 630 630 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1268 1268 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1448 1448 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1601 1601 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1751 1751 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1922 1922 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 2146 2146 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 2270 2270 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 2318 2318 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 2570 2570 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 2963 2963 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 3017 3017 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 3074 3074 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 3186 3186 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 3413 3413 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 3440 3440 N-linked (GlcNAc. . .) (Potential).
FT VARSPLIC 3430 3431 Missing (in isoform 2).
FT VARSPLIC 3430 3462 Missing (in isoform 3).
FT VARSPLIC 3430 3462 Missing (in isoform 3).
FT CONFLICT 336 336 H -> R (in Ref. 2).
FT CONFLICT 2714 2714 V -> L (in Ref. 2).
SQ SEQUENCE 3462 AA; 387525 MW; FCCF89B090B035F6 CRC64;

Query Match 95.4%; Score 596; DB 1; Length 3462;
Best Local Similarity 94.9%; Pred. No. 1.8e-54;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EOCCTIHGNAVTCCEPYGPELTCTTCTANTTASVLOFSIGSGSCRFSPSYSTTVSYAK 60
Db 231 EOCCTIHGNAVTCCEPYGPELTCTTCTANTTASVLOFSIGSGSCRFSPSYSTTVSYAK 290
Qy 61 NNTADMTQLEKTRAPSNVSTVTHILYLPKAGKGSVQFQWKQDSLRYGVEYACMAL 117
Db 291 NNTADMTQLEKTRAPSNVSTVTHILYLPKAGKGSVQFQWKQDSLRYGVEYACMAL 347

RESULT 4
ID REIN_HUMAN STANDARD; PRT; 3460 AA.
AC P78509; Q86U09; Q86U08; Q8NDV0; Q9UDQ2;
DT 26-FEB-2003 (Rel. 41, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DE 25-JAN-2005 (Rel. 46, Last annotation update)
DE Reelin precursor (EC 3.4.21.-).
GN Name=RELN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97202106; PubMed=9049633;
RA Desilva U., D'Arcangelo G., Braden V.V., Chen J., Miao G.G.,
RA Curran T., Green E.D.;
RT "The human reelin gene: isolation, sequencing, and mapping on
RT chromosome 7."
RL Genome Res. 7:157-164(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA Hillier L.M., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wyllie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Rewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali U., Andrews S., Isak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Coutney L., Kallio J.,
RA Ozerky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strommatt C.,
RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spiech J.,
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA Clifton S.W., Chisoe S.L., Marra M.A., Raymond C., Haugen E.,
RA Gillett W., Zhou Y., James R., Phelps K., Iadonoto S., Budd K.,
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA Baerbach R.A., Brent M.R., Keibler E., Filcek P., Bork P., Suyama M.,
RA Bailey J.A., Portnoy M.B., Torrents D., Chinwalla A.T., Gish W.R.,
RA Bddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA Waterston R.H., Wilson R.K.;
RT "The DNA sequence of human chromosome 7."
RL Nature 424:157-164(2003).
RN [3]
RP ALTERNATIVE SPLICING.
RX MEDLINE=99263436; PubMed=10326932; DOI=10.1006/exnr.1999.7019;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development."
RL Exp. Neurol. 156:229-238(1999).
RN [4]
RP DISEASE, AND TISSUE SPECIFICITY.
RX MEDLINE=99080080; PubMed=9861036; DOI=10.1073/pnas.95.26.15718;
RA Impagnatello F., Guidotti A.R., Pesold C., Diwedi Y., Caruncho H.,
RA Pisu M.G., Uzunov D.P., Smalheiser N.R., Davis J.M., Pandey G.N.,
RA Pappas G.D., Tuetting P., Sharma R.P., Costa E.;
RT "A decrease of reelin expression as a putative vulnerability factor in
RT schizophrenia."
RL Proc. Natl. Acad. Sci. U.S.A. 95:15718-15723(1998).
RN [5]
RP DISEASE.
RX MEDLINE=20428190; PubMed=10973257; DOI=10.1038/79246;
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahnan S.A., Grant P.E.,
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.;
RT "Autosomal recessive lissencephaly with cerebellar hypoplasia is
RT associated with human RELN mutations."
RL Nat. Genet. 26:93-96(2000).
RN [6]
RP ERATUM.
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahnan S.A., Grant P.E.,
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.;
RL Nat. Genet. 27:225-225(2001).
RN [7]
RP DISEASE.
RX MEDLINE=21217116; PubMed=11317216; DOI=10.1038/mp.4000850;
RA Persico A.M., D'Aguma L., Maiorano N., Totaro A., Militeri R.,
RA Bravaccio C., Waselink T.H., Schneider C., Melmed R., Trillo S.,
RA Montecchi F., Palermo M., Pascucci T., Pugliesi-Allegria S.,
RA Reichelt K.-L., Conciatori M., Marino R., Quattrocchi C.C., Baldi A.,
RA Zelante L., Gasparini P., Keller F.;

```


RT "Reelin gene alleles and haplotypes as a factor predisposing to
 RT autistic disorder.";
 RL Mol. Psychiatry 6:150-159 (2001).
 CC -1- FUNCTION: Extracellular matrix serine protease that plays a role
 CC in layering of neurons in the cerebral cortex and cerebellum.
 CC Regulates microtubule function in neurons and neuronal migration.
 CC Affects migration of sympathetic preganglionic neurons in the
 CC spinal cord, where it seems to act as a barrier to neuronal
 CC migration. Enzymatic activity is important for the modulation of
 CC cell adhesion. Binding to the extracellular domains of lipoprotein
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
 CC Dab1 and modulation of Tau phosphorylation (By similarity).
 CC -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P78509-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P78509-2; Sequence=VSP_005575;
 CC Name=3;
 CC IsoId=P78509-3; Sequence=VSP_005576;
 CC -1- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
 CC by the Cajal-Retzius cells and other pioneer neurons located in
 CC the telencephalic marginal zone and by granule cells of the
 CC external granular layer of the cerebellum. In adult brain,
 CC preferentially expressed in GABAergic interneurons of prefrontal
 CC cortices, temporal cortex, hippocampus and glutamatergic granule
 CC cells of cerebellum. Also expressed in fetal and adult liver.
 CC -1- DEVELOPMENTAL STAGE: Expressed in fetal and postnatal brain and
 CC liver. Expression in postnatal human brain is high in the
 CC cerebellum.
 CC -1- DOMAIN: The basic C-terminal region is essential for secretion (By
 CC similarity).
 CC -1- DISEASE: Defects in RELN are the cause of autosomal recessive
 CC lissencephaly with cerebellar hypoplasia [MIM:257320]; also known
 CC as Norman-Roberts syndrome. Some patients also displayed
 CC persistent lymphedema neonatally, and one showed accumulation of
 CC chylous or fatty, ascites fluid.
 CC -1- DISEASE: Defects in RELN may contribute to susceptibility to
 CC schizophrenia. Expression of the protein is reduced to about 50%
 CC in patients with schizophrenia.
 CC -1- DISEASE: Defects in RELN may predispose to autistic disorder. A
 CC polymorphic GGC triplet repeat located in the 5'UTR region of RELN
 CC gene, which harbors in the normal population 8 to 10 repeats, is
 CC significantly increased in autistic patients to carry 4 to 23
 CC additional repeats.
 CC -1- SIMILARITY: Belongs to the reelin family.
 CC -1- SIMILARITY: Contains 15 BNR repeats.
 CC -1- SIMILARITY: Contains 8 BGF-like domains.
 CC -1- SIMILARITY: Contains 1 reelin domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, U79716; AAC51105.1; -;
 DR EMBL, AC002067; AAM49151.1; -;
 DR EMBL, AC006981; -; NOT_ANNOTATED_CDS.
 DR EMBL, AC073208; AAP22355.1; -;
 DR EMBL, AC005101; AAP22330.1; -;
 DR EMBL, AC000121; AAB46357.2; -;
 DR EMBL, AC006316; AAD29127.1; -;
 DR EMBL, AC005064; -; NOT_ANNOTATED_CDS.
 DR HSSP; P05106; 1L5G.
 DR GeneW; HGNC:9957; RELN.
 DR MIM; 600514; -;
 DR MIM; 257320; -;

DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR002860; Glyco_hydro_BNR.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR002861; Reeler.
 DR InterPro; IPR011040; Stailase.
 DR Pfam; PR02012; BNR_15.
 DR Pfam; PR00008; EGF_2.
 DR Pfam; PR02014; Reeler_1.
 DR SMART; SM00181; EGF_5.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 6.
 DR PROSITE; PS00026; EGF_3; 5.
 DR PROSITE; PS1019; RELN_1.
 DR PROSITE; PS1019; RELN_1.
 KW Alternative splicing; Cell adhesion; Developmental protein;
 KW EGF-like domain; Extracellular matrix; Glycoprotein; Hydrolase;
 KW Lissencephaly; Repeat; Serine protease; Signal.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 3460 Reelin.
 FT DOMAIN 26 190 Reelin.
 FT DOMAIN 670 701 EGF-like 1.
 FT DOMAIN 1029 1060 EGF-like 2.
 FT DOMAIN 1408 1441 EGF-like 3.
 FT DOMAIN 1764 1795 EGF-like 4.
 FT DOMAIN 2128 2160 EGF-like 5.
 FT DOMAIN 2477 2508 EGF-like 6.
 FT DOMAIN 2852 2883 EGF-like 7.
 FT DOMAIN 3227 3259 EGF-like 8.
 FT REPEAT 592 603 BNR 1.
 FT REPEAT 798 809 BNR 2.
 FT REPEAT 951 962 BNR 3.
 FT REPEAT 1156 1167 BNR 4.
 FT REPEAT 1322 1333 BNR 5.
 FT REPEAT 1534 1545 BNR 6.
 FT REPEAT 1685 1696 BNR 7.
 FT REPEAT 1883 1894 BNR 8.
 FT REPEAT 2042 2053 BNR 9.
 FT REPEAT 2249 2260 BNR 10.
 FT REPEAT 2398 2409 BNR 11.
 FT REPEAT 2597 2608 BNR 12.
 FT REPEAT 2777 2788 BNR 13.
 FT REPEAT 2978 2989 BNR 14.
 FT REPEAT 3362 3373 BNR 15.
 FT DOMAIN 3431 3460 Arg-rich (basic).
 Query Match 91.5%; Score 572; DB 1; Length 3460;
 Best local Similarity 90.6%; Pred. No. 6.9e-52;
 Matches 106; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ECGTIMGNAVTFCEPYGPRELTTTCANTTASVLOFSIGSGCRFSYSDPSITYYAK 60
 DB 229 ECGGAIMGNNAVTFCEPYGPRELTTTGINTTASVLOFSIGSGCRFSYSDPSIIVLYAK 288
 QY 61 NNTAQMIOLEKTRASNVSTVTHILYLPREAGSGSVQPMKQDSLRVGEVYACWAL 117
 DB 289 NNSADWIOLEKTRAPSNVSTIHTILYLPDAGAGENVQPMKQENLRVGEVYACWAL 345
 RESULT 5
 ID 06Q145 PRELIMINARY; PRT; 216 AA.
 AC 06Q145;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Reelin (Fragment).
 GN Name=RELN;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Oberly E.A., Medrano J.F., Denise S.K.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY568567; AAS73245.1; -
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 216 AA; 24203 MW; 3AFDC8C22AFA3114 CRC64;
 Query Match 91.2%; Score 570; DB 2; Length 216;
 Best Local Similarity 89.7%; Pred. No. 4e-53;
 Matches 105; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ECGTGMGNAVTFCEPFGRELTTCNTTASVLOFSGSGSCRFPSDPSITVSYAK 60
 DB 72 ECGGAMHGNAVTFCEPFGRELTTCNTTASVLOFSGSGSCRFPSDPSITVSYAK 131
 QY 61 NNTADMTGLEKTPASVSVTHILVYPERAKGSVQFQKQDSLNGEYEAACWAL 117
 DB 132 NNTADMTGLEKTPASVSVTHILVYPERAKGSVQFQKQDSLNGEYEAACWAL 188
 RESULT 6
 REIN_CHICK STANDARD; PRT; 3209 AA.
 ID REIN_CHICK
 AC 093574;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Reelin (EC 3.4.21.-) (Fragment).
 GN Name=RELN;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20320899; PubMed=10861519;
 RX DOI=10.1002/1096-9861(20000703)42:3<448::AID-CNE10>3.3.CO;2-W;
 RA Bernier B., Bar I., D'Arcangelo G., Curran T., Goffinet A.M.;
 RT "Reelin mRNA expression during embryonic brain development in the chick.";
 RT J. Comp. Neurol. 422:448-463(2000).
 CC -1- FUNCTION: Extracellular matrix serine protease that plays a role in layering of neurons in the cerebral cortex and cerebellum. Regulates microtubule function in neurons and neuronal migration. Affects migration of sympathetic preganglionic neurons in the spinal cord, where it seems to act as a barrier to neuronal migration. Enzymatic activity is important for the modulation of cell adhesion. Binding to the extracellular domains of lipoprotein receptors VLDLR and ApoER2 induces tyrosine phosphorylation of Dab1 and modulation of tau phosphorylation (by similarity).
 CC -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (by similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -1- DOMAIN: The basic C-terminal region is essential for secretion (by similarity).
 CC -1- SIMILARITY: Belongs to the reelin family.
 CC -1- SIMILARITY: Contains 15 BNR repeats.
 CC -1- SIMILARITY: Contains 8 EGF-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; AF090441; AAC3559.1; -
 DR HSSP; P05106; 1JV2.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR002860; Glyco_hydro_BNR.
 DR InterPro; IPR011040; Stalfase.
 DR Pfam; PF02012; BNR; 15.
 DR Pfam; PF00008; EGF; 3.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 7.
 DR PROSITE; PS00026; EGF_3; 3.
 KW Cell adhesion; Developmental protein; EGF-like domain; Glycoprotein; Hydroxylase; Matrix protein; Repeat; Serine protease.
 FT NON TER 1
 FT DOMAIN 1 1
 FT DOMAIN 418 449 EGF-like 1.
 FT DOMAIN 777 808 EGF-like 2.
 FT DOMAIN 1157 1190 EGF-like 3.
 FT DOMAIN 1513 1544 EGF-like 4.
 FT DOMAIN 1877 1909 EGF-like 5.
 FT DOMAIN 2226 2257 EGF-like 6.
 FT DOMAIN 2601 2632 EGF-like 7.
 FT DOMAIN 2976 3008 EGF-like 8.
 FT DOMAIN 340 351 BNR 1.
 FT REPEAT 546 557 BNR 2.
 FT REPEAT 699 710 BNR 3.
 FT REPEAT 904 915 BNR 4.
 FT REPEAT 1070 1081 BNR 5.
 FT REPEAT 1283 1294 BNR 6.
 FT REPEAT 1434 1445 BNR 7.
 FT REPEAT 1632 1643 BNR 8.
 FT REPEAT 1791 1802 BNR 9.
 FT REPEAT 1998 2009 BNR 10.
 FT REPEAT 2147 2158 BNR 11.
 FT REPEAT 2346 2357 BNR 12.
 FT REPEAT 2526 2537 BNR 13.
 FT REPEAT 2727 2738 BNR 14.
 FT REPEAT 3111 3122 BNR 15.
 FT DOMAIN 3180 3209 Arg-rich (basic).
 FT CARBOHYD 5 5 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 37 37 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 53 53 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 376 376 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1014 1014 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1195 1195 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1348 1348 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1669 1669 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1893 1893 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2017 2017 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2065 2065 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2317 2317 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2710 2710 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2764 2764 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2821 2821 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2933 2933 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 3160 3160 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 3187 3187 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 3209 AA; 361288 MW; 81A7B676BCAA3D1 CRC64;
 Query Match 66.9%; Score 418; DB 1; Length 3209;
 Best Local Similarity 83.9%; Pred. No. 2.5e-35;
 Matches 78; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
 QY 25 TTCLNTTASVLOFSGSGSCRFPSDPSITVSYAKNNTADMTGLEKTPASVSVTHI 84
 DB 1 TTCLNTTASVLOFSGSGSCRFPSDPSITVSYKNSADMTGLEKTPASVSVTHI 60
 QY 85 LYLPERAKGSVQFQKQDSLNGEYEAACWAL 117
 DB 61 LYLPERAKGSVQFQKQDSLNGEYEAACWAL 93
 RESULT 7
 ID 08C978 PRELIMINARY; PRT; 279 AA.
 AC 08C978;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (T-EmbRel. 26, last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:A730023J04 product:reelin, full insert
DE sequence.
GN Name=ReIn;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning."
RT Meth. Enzymol. 303:19-44(1999).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RT Nature 409:685-690(2001).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC THE FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.153600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multiplexed capillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Horii F., Imoutani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohashi N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaki-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK042773; BAC31362.1; -
DR MGD; MGI:103023; ReIn.
GO GO:0005578; C.eintracellilar matrix (sensu Metazoa); TNS.
GO GO:0005615; C.eintracellilar space; IDA.

```

DR GO: GO:0004252; F:serine-type endopeptidase activity; IMP.
DR GO: GO:0007420; P:brain development; IMP.
DR GO: GO:0007420; P:brain development; IMP.
DR GO: GO:0016477; P:cell migration; IMP.
DR GO: GO:0009044; P:cellular morphogenesis during differentiation; IMP.
DR GO: GO:0007529; P:establishment of synaptic specificity at ne. . .; IMP.
DR GO: GO:0010001; P:glial cell differentiation; IMP.
DR GO: GO:00051057; P:positive regulation of small GTPase mediate. . .; IDA.
DR InterPro: IPRO002861; Reeler.
DR Pfam: PF02014; Reeler; 1.
DR Sequence 279 AA; 30183 MW; 26741D24A4F8EC97 CRC64;

Query Match 35.7%; Score 223; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ECGTTHGNAVTFCPEYGPRLTTTCLNTTASVLOPFSIG 41
Db 230 ECGTTHGNAVTFCPEYGPRLTTTCLNTTASVLOPFSIG 270

RESULT 8
RELN_BOVIN STANDARD; PRT; 52 AA.
AC OGN117;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2004 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Reelin (EC 3.4.21.-) (Fragment).
GN Name=RELN;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RF Sequence from N.A.
RC STRAIN=Brown Swiss;
RA Speidel S.E., Oberly E.A., Ben Abdallah M., Denise S.K.;
RT "Genetic analysis of candidate gene (RELN) for Weaver Syndrome in
   Brown Swiss cattle."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Extracellular matrix serine protease that plays a role
   in layering of neurons in the cerebral cortex and cerebellum.
   Regulates microtubule function in neurons and neuronal migration.
   Affects migration of sympathetic preganglionic neurons in the
   spinal cord, where it seems to act as a barrier to neuronal
   migration. Enzymatic activity is important for the modulation of
   cell adhesion. Binding to the extracellular domains of lipoprotein
   receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
   Dab1 and modulation of tau phosphorylation (By similarity).
CC -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
   similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
   by the Cajal-Retzius cells and other pioneer neurons located in
   the telencephalic marginal zone and by granule cells of the
   external granular layer of the cerebellum.
CC -1- SIMILARITY: Belongs to the reelin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: AF323904; AAF64286.1; -
KW Cell adhesion; Developmental protein; Hydrolase; Matrix protein;
KM Serine protease.
FT NON_TER 1 1
FT TER 52 52
SQ SEQUENCE 52 AA; 6069 MW; B4699D90CDC998F7 CRC64;

```



```
RESULT 12
O6PX75 PRELIMINARY; PRT; 205 AA.
AC O6PX75;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Reelin (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylia; Crocodyline; Crocodylus.
OX NCBI_TaxId=8501;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=2428833; PubMed=12541309; DOI=10.1002/cme.10573;
RA Tislar F., Lambert D., Rouvolet C., Sire J.Y., Meyer G., Goffinet A.M.;
RT "Reelin expression during embryonic brain development in Crocodylus
RT niloticus".
RL J. Comp. Neurol. 457:250-262 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Tislar F., Goffinet A.M.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY572417; AAS78666.1; -
DR InterPro; IPR008976; PLAT_LH2.
FT NON_TER 1
SQ SEQUENCE 205 AA; 23709 MW; A09163964AC9A88 CRC64;

Query Match 14.2%; Score 88.5; DB 2; Length 205;
Best Local Similarity 29.1%; Pred. No. 0.29; Mismatches 42; Indels 21; Gaps 8;
Matches 34; Conservative 20;

OY 3 CGTI---MHGNAVTF--CEPYGPRELTTCLNTTASVLOF--SIGS-----GSCRFYSYD 51
DB 42 CQGLAPYAHGDSLRYNGCQ---IROAITPELDLITRASKIMFVLOIGSISQTSCTNLS 98
OY 52 P-----SITVSYAKNNTADWIOLEKIRAPSNVSVIHLILY-LPEAKGESVQFOKQ 102
DB 99 PNTVDKAVLQYSVNNGITW-QVIAQHQPDPFIOQRVSYNVPLBARMKGVLLRMWQ 154

RESULT 13
O8AYTO PRELIMINARY; PRT; 308 AA.
AC O8AYTO;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Reelin (Fragment).
GN Name=Reelin; Synonym=ReIn;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikkawa S., Terashima T.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB072424; BAC20286.1; -
DR ZFIN; ZDB-GENE-040427-1; reln1.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR008976; PLAT_LH2.
DR SMART; SM00181; EGF_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 1.
FT NON_TER 1
SQ SEQUENCE 308 AA; 33858 MW; B2719511CC04C2F CRC64;
```

```
Query Match 14.0%; Score 87.5; DB 2; Length 308;
Best Local Similarity 27.4%; Pred. No. 0.61;
Matches 32; Conservative 19; Mismatches 45; Indels 21; Gaps 8;

OY 3 CGTI---MHGNAVTF--CEPYGPRELTTCLNTTASVLOF--SIGS-----GSCRFYSYD 51
DB 172 CQGLAPYAHGDSLRYNGCQ---IROAITPELDLITRASKIMFVLOIGSVSQTDSCTNLS 228
OY 52 P-----SITVSYAKNNTADWIOLEKIRAPSNVSVIHLILY-LPEAKGESVQFOKQ 102
DB 229 PNTVDKAVLQYSVNNGSMHVIQ-HQPKDITKQKRSYNPLBARMKGVLLRMWQ 284

RESULT 14
O93575 PRELIMINARY; PRT; 615 AA.
AC O93575;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Reelin (Fragment).
OS Emyx orbicularis (European pond turtle).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Cryptodira; Testudinoidea; Emydidae; Emyx.
OX NCBI_TaxId=82168;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernier B., Goffinet A.M.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: Contains 2 EGF-like domains.
DR EMBL; AF090843; AAC35993.1; -
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002860; Glyco_hydro_BNR.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR01040; Sialidase.
DR Pfam; PF02012; BNR_2.
DR Pfam; PF00008; EGF_1.
DR SMART; SM00181; EGF_2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 615 AA; 69837 MW; A074D7AD9979C9DB CRC64;

Query Match 14.0%; Score 87.5; DB 2; Length 615;
Best Local Similarity 29.1%; Pred. No. 1.4;
Matches 34; Conservative 20; Mismatches 42; Indels 21; Gaps 8;

OY 3 CGTI---MHGNAVTF--CEPYGPRELTTCLNTTASVLOF--SIGS-----GSCRFYSYD 51
DB 450 CQGLAPYAHGDSLRYNGCQ---IROAITPELDLITRASKIMFVLOIGSISQTSCTNLS 506
OY 52 P-----SITVSYAKNNTADWIOLEKIRAPSNVSVIHLILY-LPEAKGESVQFOKQ 102
DB 507 PNTVDKAVLQYSVNNGITW-QVIAQHQPDPFIOQRVSYNVPLBARMKGVLLRMWQ 562

RESULT 15
O870G8 PRELIMINARY; PRT; 639 AA.
AC O870G8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glucoamyrase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxId=5061;
RN [1]
```

RP SEQUENCE FROM N.A.
 RC STRAIN=T21;
 RX MEDLINE=95066018; PubMed=7975554;
 RA Zhong L., Tang G., Yang K.;
 RT "Isolation and sequencing of glucoamylase gene from a glucoamylase
 RT over producing strain."
 RL Wei Sheng Wu Xue Bao 34:184-190 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T21;
 RA Zhong L., Qiao D., Tang G., Yang K.;
 RT "Cloning, sequencing and comparison of the 5' flanking regions of *gl*a
 RT gene from high and low glucoamylase-producing strains of *Aspergillus*
 RT *niger*."
 RL Wei Sheng Wu Xue Bao 36:181-186 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T21;
 RA Zhong L., Qiao D., Tang G., Yang K.;
 RL Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY250996; AAP04499.1; --
 DR HSSP; P04064; IKM.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO; GO:0000272; P:polysaccharide catabolism; IEA.
 DR InterPro; IPR008291; Glu-a-glcSD_SBD.
 DR InterPro; IPR00165; Glyco_hydro_15.
 DR InterPro; IPR002044; Glyco_hydro_CBD.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CBM_20; 1.
 DR Pfam; PIRSF01031; Glu-a-glcSD_SBD; 1.
 DR PRINTS; PR00736; GUYHRLAS515.
 DR ProDom; PD001568; Glyco_hydro_CBD; 1.
 QO SEQUENCE 639 AA; 68308 MW; 025AD856B6542B69 CRC64;

Query Match	13.0%	Score 81	DB 2	Length 639
Best Local Similarity	25.2%	Pred. No. 7.4		
Matches 34	Conservative 15	Mismatches 54	Indels 32	Gaps 4

```

Qy      9  GNAATVFCPEYGRRELTTTCAMTTTASVLQFSGSGSCRFYS-----DPSITVSYAKN-- 61
Db      496 GGTITTAATPTSGSVSTSTSTKTATNAKSKTSTSTSTCTTPPAVAATFDLTATTTGGINIY 555
Qy      62  -----NTADWIOLEKLRAPSNVSTVIHILYLPBEAKSGSVQF-----Q 99
Db      556 LVGSIQQLGDMETISGIALSADKYTSSDPLMTVTYTLPR--AGSEFEYKFIKRIESDSVE 612
Qy      100 WKQDSLRYGGEVYBEAC 114
Db      613 WESDPNRRYTVQAC 627

```

Search completed: August 4, 2005, 15:35:29
Job time : 170 secs

This Page Blank (uspto)